

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:07:08 ; Search time 113.31 Seconds  
(without alignments)  
9.081 Million cell updates/sec

Title: US-09-165-546A-10  
Perfect score: 92  
Sequence: 1 PLPVPVLLKERTVSGNI 18

Scoring table: BLOSUMP2  
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 165078

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0401.\*  
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	81.5	88.6	19	21	Y52437 Human tumour antig
2	52	56.5	10	20	Y05998 Human cancer antig
3	45	48.9	9	20	Y06026 Human cancer antig
4	43	46.7	9	20	Y06027 Human cancer antig
5	41.5	45.1	17	21	Y52436 Human tumour antig
6	37	40.2	9	16	R82133 Melanoma-specific
7	37	40.2	23	20	W97452 Antigenic site of
8	35	38.0	9	21	Y79753 NY-ESO-1 derived p
9	35	38.0	10	20	Y06001 Human cancer antig
10	33	35.9	9	16	R82132 Melanoma-specific
11	33	35.9	10	16	R82174 Melanoma-specific

12	32	34.8	15	17	W05451	SH3-binding peptid
13	32	34.8	15	18	W37659	peptide motif contai
14	32	34.8	15	18	W38924	Reagent of GTP-bin
15	32	34.8	20	11	R04057	Human p160 interna
16	32	34.8	20	18	W31198	Human peptide hkv1
17	32	34.8	22	20	Y41645	Antigenic site of
18	32	34.8	23	20	W97453	Mammalian ion chan
19	32	34.8	25	20	Y41612	Fragment of human
20	32	34.8	25	20	Y38419	Human CERB2 oncoge
21	31	33.7	9	15	Y37964	Immunogenic peptid
22	31	33.7	9	15	Y37970	Immunogenic peptid
23	31	33.7	9	20	Y45534	Immunogenic peptid
24	31	33.7	9	20	Y45540	Immunogenic peptid
25	31	33.7	10	15	Y37993	Human CERB2 oncoge
26	31	33.7	10	20	Y45561	Immunogenic peptid
27	31	33.7	11	20	Y45563	Immunogenic peptid
28	31	33.7	11	20	Y41628	Mammalian ion chan
29	31	33.7	15	18	W26629	Signalling inosito
30	31	33.7	15	21	Y93064	Transforming growt
31	31	33.7	16	14	R30832	Sequence of peptid
32	31	33.7	21	19	W80796	Fragment of human
33	31	33.7	21	20	Y36587	Human DPI (HDP1) i
34	31	33.7	23	20	Y38439	Human secreted pro
35	30	32.6	24	21	Y87770	Potato THT protein
36	30	32.6	13	17	W05482	SH3-binding peptid
37	30	32.6	13	17	W05415	Src SH3 domain-bin
38	30	32.6	13	17	W11101	Src SH3 domain-bin
39	30	32.6	13	18	W25514	SH3 synthetic bind
40	30	32.6	15	16	R65210	Murine atromalin-1
41	30	32.6	17	17	W05421	SH3 domain peptid
42	30	32.6	17	21	Y66032	FDF-5 mutant pepti
43	30	32.6	18	21	Y66033	FDF-5 mutant pepti
44	30	32.6	25	17	R99000	Human cytomagalovi
45	30	32.6	25	18	W31702	Human cytomagalovi

## ALIGNMENTS

## RESULT 1

Y52437  
ID Y52437 standard; Protein; 19 AA.

XX Y52437;

XX  
DT 15-FEB-2000 (first entry)

XX Human tumour antigen NY-ESO-1 peptide #10.

XX Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
KW T-cell; helper; stimulation; proliferation; treatment;  
KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
KW lymphoma.

XX Synthetic.

OS Homo sapiens.

XX W09953938-Al.

XX 28-OCT-1999.

XX 24-MAR-1999; 99WO-US06875.

XX 17-APR-1998; 98US-0062422.

XX 02-OCT-1998; 98US-0165546.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;

PI Gure A, Ritter G;

XX WPI; 2000-038483/03.

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
PT useful for therapeutic and diagnostic purposes -  
PS Claim 4; Page 22; 49pp; English.  
XX Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
CC localisation studies revealed it to be expressed at high levels  
CC in normal ovary and testis but not in normal colon, kidney, liver,  
CC brain, oesophagus and skin. It was expressed in certain tumours and  
CC tumour cell lines with some degree of frequency - these included  
CC melanoma specimens and cell lines, and breast and bladder cancer  
CC specimens, with expression in other tumour types being sporadic.  
CC These NY-ESO-1-derived peptides may be used in methods and  
CC compositions used for the treatment, diagnosis and prevention of  
CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
CC or lymphoma) and to stimulate the proliferation of T cells.  
XX  
SQ Sequence 19 AA;

Query Match 88.6%; Score 81.5; DB 21; Length 19;  
Best Local Similarity 94.7%; Pred. No. 1.le-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 PLPVPVGLL-KEFTVSGNI 18  
Db 1 plpvpvllkkeftvsgni 19  
|||||||

RESULT 2  
ID Y05998 standard; Peptide; 10 AA.  
XX  
AC Y05998;  
XX  
XX  
DT 16-AUG-1999 (first entry)  
DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.  
XX

KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine; human leukocyte antigen; HLA.  
XX  
OS Homo sapiens.  
XX  
XX  
PN W09918206-A2.  
XX  
PD 15-APR-1999.  
XX  
XX  
XX 21-SEP-1998; 98WO-US19609.  
XX  
XX 08-OCT-1997; 97US-0061428.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Rosenber SA, Wang RF;  
XX  
XX WPI; 1999-277270/23.  
XX  
XX Cancer antigen NY ESO1/CAG-3  
XX  
XX Example 10; Page 42; 88pp; English.

CC This peptide was identified as an HLA peptide motif following a  
CC screen for epitopes from the coding region of human ESO-1/CAG-3  
CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
CC The present peptide (ranked 11) corresponds to amino acid residues  
CC 115-124 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
CC tumour antigen capable of eliciting an antigen specific immune  
CC response by T cells. Cancer peptides (see Y05967-87) derived from  
CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
CC vaccines. A claimed method of preventing or inhibiting cancer  
CC involves administering a cancer peptide, with or without an HLA  
CC molecule. The cancer peptides form part of, or are derived  
CC from, cancers such as primary or metastatic melanoma, thymoma,  
CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
CC cancer, cervical cancer, bladder cancer, kidney cancer and  
CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
CC thyroid cancers.  
XX  
SQ Sequence 10 AA;

Query Match 56.5%; Score 52; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0074;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPVPVGLLK 10  
Db 1 plpvpvllk 10  
|||||||

RESULT 3  
ID Y06026 standard; Peptide; 9 AA.  
XX  
AC Y06026;  
XX  
XX 16-AUG-1999 (first entry)  
DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.  
XX

KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine; human leukocyte antigen; HLA.  
XX  
OS Homo sapiens.  
XX  
XX  
PN W09918206-A2.  
XX  
PD 15-APR-1999.  
XX  
XX 21-SEP-1998; 98WO-US19609.  
XX  
XX 08-OCT-1997; 97US-0061428.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Rosenber SA, Wang RF;  
XX  
XX WPI; 1999-277270/23.  
XX  
XX Cancer antigen NY ESO1/CAG-3  
XX  
XX Example 10; Page 43; 88pp; English.

CC This peptide was identified as an HLA peptide motif following a  
CC screen for epitopes from the coding region of human ESO-1/CAG-3  
CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
CC The present peptide (ranked 9) corresponds to amino acid residues  
CC 116-124 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
CC tumour antigen capable of eliciting an antigen specific immune

CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer, and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 SQ Sequence 9 AA;

Query Match 48.9%; Score 45; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPVPGVLEK 10  
 Db 1 lvpvgvllk 9

RESULT 4  
 Y06027  
 ID Y06027 standard; Peptide; 9 AA.

XX  
 AC Y06027;  
 XX  
 DT 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

XX WO9918206-A2.

XX  
 PD 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Rosenberg SA, Wang RF;

XX WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 10) corresponds to amino acid residues  
 CC 120-128 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,

CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 SQ Sequence 9 AA;

Query Match 46.7%; Score 43; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GVLLKFTV 14  
 Db 1 gvlkftv 9

RESULT 5  
 Y52436  
 ID Y52436 standard; Protein; 17 AA.

XX Y52436;

XX 15-FEB-2000 (first entry)

XX Human tumour antigen NY-ESO-1 peptide #9.

KW Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
 KW T-cell; helper; stimulation; proliferation; treatment;  
 KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
 KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
 KW lymphoma.

XX Synthetic.

XX Homo sapiens.

XX WO9953938-A1.

XX 28-OCT-1999.

XX 24-MAR-1999; 99WO-US06875.

XX 17-APR-1998; 98US-0062422.

XX 02-OCT-1998; 98US-0165546.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
 XX Gure A, Ritter G;

XX WPI; 2000-038483/03.

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
 XX useful for therapeutic and diagnostic purposes

XX Claim 4; Page 22; 49pp; English.

XX Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels  
 CC in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and  
 CC tumour cell lines with some degree of frequency - these included  
 CC melanoma specimens and cell lines, and breast and bladder cancer  
 CC specimens, with expression in other tumour types being sporadic.  
 CC These NY-ESO-1-derived peptides may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of  
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
 CC or lymphoma) and to stimulate the proliferation of T cells.

XX  
SQ Sequence 17 AA;

Query Match 45.1%; Score 41.5; DB 21; Length 17;  
Best Local Similarity 91.7%; Pred. No. 0.96;  
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7 VLLKEFTVSGNI 18  
| | | | | | | | | |  
Db 1 vllkef-vsgni 11

RESULT 6

R82133  
ID R82133 standard; peptide; 9 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX  
AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

W97452;

19-MAY-1999 (first entry)

Antigenic site of HN protein loop beta-1L01.

Antigenic site; haemagglutinin-neuraminidase; HN; paramyxoviridae;  
virus epitope; attachment protein; vaccine; immunodominant epitope.

Bovine parainfluenza virus.

WO9902695-A2.

21-JAN-1999.

08-JUL-1998; 98WO-NL00390.

08-JUL-1997; 97EP-0202100.

(DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.

Langedijk JPM, Van Oirschot JT;

WPI; 1999-120896/10.

Isolated proteinaceous substance - comprising at least one virus  
epitope derived from an attachment protein of a paramyxovirus

Disclosure; Page 41; 63pp; English.

W97452-571 represent antigenic sites derived from the  
haemagglutinin-neuraminidase (HN) protein of the paramyxoviridae.  
The specification describes 3-D models identifying a proteinaceous  
substance comprising at least one virus epitope derived from the  
attachment protein, which corresponds to an antigenic site present on  
one of the loops of HN. The antigenic sites can be used to produce  
vaccines, to detect the viruses, and to select the immunodominant  
epitope.

Sequence 23 AA;

Query Match 40.2%; Score 37; DB 20; Length 23;  
Best Local Similarity 47.1%; Pred. No. 8.4;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LPVPGVLLKEFTVSGNI 18

Db 6 ipgpgllatsttngci 22

RESULT 8

Y79753

ID Y79753 standard; Peptide; 9 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Query Match 40.2%; Score 37; DB 16; Length 9;  
Best Local Similarity 75.0%; Pred. No. 3.2e+05;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPVPGVLL 9

Db 2 mpvpgill 9

RESULT 7

W97452

ID W97452 standard; peptide; 23 AA.

25-JUN-1999; 99WO-US14493.

06-JAN-2000.

WO200000824-A1.

Homo sapiens.

Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;

HLA binding; human leukocyte antigen; cytolytic T cell; CTL;

cytostatic; melanoma; synovial sarcoma.

XX 26-JUN-1998; 98US-0105839.  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 XX Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;  
 PI Chen Y, Gure A, Old LJ;  
 XX WPI; 2000-170933/15.  
 XX Determining the possible presence of breast, endometrial, colorectal,  
 PT lung, bladder or head-neck cancer  
 XX  
 PS Example 13; Page 26; 40pp; English.  
 XX  
 CC A method has been developed for determining the possible presence of a  
 CC cancer, which is not melanoma or synovial sarcoma. The method comprises  
 CC assaying a sample taken from the subject to determine the expression of  
 CC an SSX gene, and determining the expression as a determination of the  
 CC possible presence of cancer. Expression of SSX1 gene indicates possible  
 CC presence of breast, endometrial, colorectal, lung, bladder or head-neck  
 CC cancer. These cancers are also detected by SSX2 and SSX4 gene expression.  
 CC SSX2 gene expression additionally indicates possible presence of  
 CC lymphoma, renal cell cancer, glioma and prostate cancer. Expression of  
 CC SSX4 gene also indicates possible presence of ovarian or stomach cancer.  
 CC SSX5 gene expression indicates the same cancers as SSX1, except breast  
 CC cancer. Determining expression of SSX gene can be used to monitor  
 CC progress of melanoma or synovial sarcoma, which is not cancer. The  
 CC SSX-derived peptide complex stimulates proliferation of cytolytic T  
 CC cells. This is useful for treating cancer, especially melanoma. Y78464  
 CC to Y78468 represent specifically claimed HLA binding peptides for use in  
 CC the method of the invention. Y88452 to Y88465 represent PCR primers used  
 CC in the isolation of SSX genes in the exemplification of the present  
 CC invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides  
 CC derived from SSX proteins or NY-ESO-1, which are used in the  
 CC exemplification of the present invention.  
 XX  
 XX Sequence 9 AA;

Query Match 38.0%; Score 35; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FTVSGNI 18  
 Db 1 ftvsgni 7

RESULT 9  
 Y06001  
 ID Y06001 standard; Peptide; 10 AA.  
 XX  
 AC Y06001;  
 XX

DT 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

OS Homo sapiens.

XX W09918206-A2.

PN 15-APR-1999.

XX

PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Rosenberg SA, Wang RF;  
 XX  
 DR WPI; 1999-277270/23.  
 XX  
 XX Cancer antigen NY ESO1/CAG-3  
 PT  
 PS Example 10; Page 42; 88pp; English.  
 XX

CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human NY ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 14) corresponds to amino acid residues  
 CC 126-135 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 10 AA;

Query Match 38.0%; Score 35; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FTVSGNI 18  
 Db 1 ftvsgni 7

RESULT 10

R82132  
 ID R82132 standard; peptide; 9 AA.  
 XX  
 AC R82132;

DT 25-MAR-1996 (first entry)

DE Melanoma-specific mutant immunogen epitope 9mer peptide.

XX Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW cytotoxic T cell; lymphocyte; HLA-A2.

OS Homo sapiens.

XX W09522561-A2.

PD 24-AUG-1995.

PF 16-FEB-1995; 95WO-US01991.

XX 29-APR-1994; 94US-0234784.

PR 16-FEB-1994; 94US-0197399.

XX (UYVI-) UNIV VIRGINIA PATENT FOUND.

XX Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;

XX WPI; 1995-302688/39.

XX

PT Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pMel-17 - are highly potent stimulators of HLA-A2+CTL's useful in  
 PT adoptive immuno-therapy

PS Example 8; Page 51; 148pp; English.

XX A melanoma-specific immunogen homologous with pMel-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82194 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMel-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.

XX Sequence 9 AA;

Query Match 35.9%; Score 33; DB 16; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 3.2e+05;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPVPGVL 8  
 :|||:|  
 Db 3 mpvpgil 9

RESULT 11

R82174 ID R82174 standard; peptide; 10 AA.

XX R82174;

XX 25-MAR-1996 (first entry)

XX Melanoma-specific mutant immunogen epitope 10mer peptide.

XX Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW cytotoxic T cell; lymphocyte; HLA-A2.

XX Homo sapiens.

XX WO9522561-A2.

XX 24-AUG-1995.

XX 16-FEB-1995; 95WO-US01991.

XX 29-APR-1994; 94US-0234784.

XX 16-FEB-1994; 94US-0197399.

XX (UYVI-) UNIV VIRGINIA PATENT FOUND.

XX Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;

XX WPI; 1995-302688/39.

XX Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pMel-17 - are highly potent stimulators of HLA-A2+CTL's useful in  
 PT adoptive immuno-therapy

PS Example 8; Page 52; 148pp; English.

XX A melanoma-specific immunogen homologous with pMel-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82194 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMel-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.

XX Sequence 10 AA;

Query Match 35.9%; Score 33; DB 16; Length 10;  
 Best Local Similarity 71.4%; Pred. No. 16;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPVPGVL 8  
 :|||:|  
 Db 4 mpvpgil 10

RESULT 12

W05451 ID W05451 standard; Peptide; 15 AA.

XX W05451;

XX 24-FEB-1998 (first entry)

XX SH3-binding peptide bSH3002.

XX Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;  
 KW cellular signalling element; cellular structural element; malignancy;  
 KW protein identification; functional domain; protein screening;  
 KW cellular signal transduction process; binding peptide.

XX Synthetic.

XX W09631625-A1.

XX 10-OCT-1996.

XX 04-APR-1996; 96WO-US04454.

XX 03-APR-1996; 96US-0630915.

XX 07-APR-1995; 95US-0417872.

XX (CYTO-) CYTOGEN CORP.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Fowikes DM, Hoffman N, Kay BK, McConnell SJ, Sparks AB;

XX WPI; 1996-465045/46.

XX Identifying polypeptide(s) having specific functional domain (esp.  
 PT SH3 domain) - comprises detecting selective binding to recognition  
 PT unit, regardless of sequence homology

PS Example; Fig 12A; 174pp; English.

XX W05445-W05492 represent Src-homology region 3 (SH2) domain binding  
 CC peptides. These sequences were used as parts of multivalent recognition  
 CC unit complexes used in the method of the invention. The method of the  
 CC invention is for identifying polypeptides containing functional domains  
 CC of interest (especially SH3 domains). It comprises contacting a  
 CC multivalent recognition unit (RU) complex with a number of peptides and  
 CC identifying polypeptides having a selective binding affinity for the RU  
 CC complex. The method is based on functional similarities and does not rely  
 CC on sequence similarities. Prior methods only gave limited success for  
 CC identifying proteins containing an SH3 domain due to the minimal sequence  
 CC homology among known SH3 proteins. Multivalent RU complexes are  
 CC particularly suited to screening for polypeptides containing functional  
 CC domains that are similar to, but not identical in sequence to, the  
 CC original target functional domain. The new method enables proteins having  
 CC a common function to be identified. Identification of novel SH3 proteins  
 CC will be useful for a better understanding of cell growth, malignancy,  
 CC signal transduction processes, etc. New candidate drugs can be  
 CC identified, and their specificities (e.g. pharmacological activities) can  
 CC be assessed using the method of the invention.

XX Sequence 15 AA;

Query Match 34.8%; Score 32; DB 17; Length 15;  
 Best Local Similarity 85.7%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVPGV 7  
 III III  
 DB 9 plpdpgv 15

RESULT 13  
 W37659  
 ID W37659 standard; Peptide; 15 AA.  
 XX  
 AC  
 XX W37659;  
 DT 23-APR-1998 (first entry)  
 DE PPPPV motif containing peptide BSH3002 used to bind WW domains.  
 XX  
 KW Peptide recognition unit; WW domain; cell signalling; growth regulation;  
 KW cytoskeleton organisation; targeted drug screening; modulator;  
 KW WW domain interaction; YAP protein; dystrophin.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9737223-A1.  
 XX  
 PD 09-OCT-1997.  
 XX  
 PF 03-APR-1997; 97WO-US05547.  
 XX  
 PR 03-APR-1996; 96US-0630916.  
 XX  
 PA (CYTO-) CYTOGEN CORP.  
 PA (UYNC-) UNIV NORTH CAROLINA.  
 XX  
 PI Fowlkes DM, Kay BK, Pirozzi G;  
 XX WPI; 1997-503234/46.  
 DR  
 XX  
 PT Identifying cell signalling and growth regulatory polypeptides by  
 PT reaction with multivalent recognition complex - polypeptides are  
 PT useful in targeted drug selection  
 XX  
 PS Example 6.3; Fig 7; 220pp; English.

Peptides W37653-77 contain PPPPV-like motifs. The PPPV motif is found in the proline rich regions of WBP-1 and WBP-2 proteins. Peptides containing this residue have been shown to bind the YAP WW domain, but not the WW domain from dystrophin or to a panel of SH3 domains. Peptides W37653-77 were biotinylated and complexed with alkaline streptavidin, and used in a cross affinity mapping experiment. They were tested for their ability to bind to the 12 individual novel WW domains of WWP1 (W36794), WWP2 (W36795), WWP3 (W37696) and WWP4 (W36797), which were expressed as glutathione-S-transferase expression proteins. The present peptide, derived from a potassium channel, does not bind to WW domains of the novel proteins. The WW domain is a small functional domain. Its name is derived from the observation that two tryptophan residues, one in the amino terminal portion of the WW domain and one in the carboxyl terminal portion, are conserved. Most proteins containing WW domains have a function involving cell signalling and growth regulation or the organisation of the cytoskeleton. Polypeptides containing a WW domain are identified by treating a multivalent recognition unit complex that has selective binding affinity for a WW domain, with many polypeptides and identifying those with selective affinity for the complex. Proteins containing WW domains are used for targeted drug screening, i.e. to identify potential modulators of specific WW domain interactions.

QY 1 PLPVPGV 7  
 III III  
 DB 9 plpdpgv 15

Query Match 34.8%; Score 32; DB 18; Length 15;  
 Best Local Similarity 85.7%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVPGV 7  
 III III  
 DB 9 plpdpgv 15

RESULT 14  
 W38924  
 ID W38924 standard; peptide; 15 AA.  
 XX  
 AC W38924;  
 XX  
 DT 27-MAR-1998 (first entry)  
 DE Peptide resembling an SH3 domain binding peptide SEQ ID NO:321.  
 XX  
 KW Cortactin; SH3 domain; binding peptide; Src homology region 3;  
 KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck;  
 KW Abl; PLCgamma; p53bp2; Crk; Yes; Grb2.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9730074-A1.  
 XX  
 PD 21-AUG-1997.  
 XX  
 PF 14-FEB-1997; 97WO-US02298.  
 XX  
 PR 16-FEB-1996; 96US-0602999.  
 XX  
 PA (CYTO-) CYTOGEN CORP.  
 PA (UYNC-) UNIV NORTH CAROLINA.  
 XX  
 PI Der CJ, Fowlkes DM, Kay BK, Quilliam LA, Rider JE;  
 PI Sparks AB, Thorn JM;  
 XX WPI; 1997-424972/39.  
 DR  
 XX  
 PT Src homology region 3 binding peptide - used to activate Src  
 PT tyrosine kinase(s) and to stimulate immune response by increasing  
 PT production of certain lymphokine(s), e.g. interleukin-1  
 XX  
 PS Claim 22; Page 90; 131pp; English.

The present sequence represents a peptide which resembles a Src homology region 3 (SH3) binding peptide. SH3 binding peptides are selected from: (a) peptides which bind the SH3 domain of Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c) peptides which bind the SH3 domain of Abl; (d) peptides which bind the SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3 domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain of Grb2. The purified binding peptides can be used in the method to identify inhibitors of their binding to their respective SH3 domains, which could be used to modulate the pharmacological activity of proteins or polypeptide containing the SH3 domain. The peptides can also be used to activate Src or Src-related protein tyrosine kinases, to stimulate the immune response by increasing the production of certain lymphokines, e.g. tumour necrosis factor-alpha and interleukin-1, or to deliver a conjugated molecule to certain cellular compartments containing Src or Src related proteins.

QY 1 PLPVPGV 7  
 III III

Query Match 34.8%; Score 32; DB 18; Length 15;  
 Best Local Similarity 85.7%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVPGV 7  
 III III

Db 9 plpdpv 15

## RESULT 15

R04057  
ID R04057 standard; peptide; 20 AA.XX  
AC R04057;XX  
DT 17-MAR-1993 (first entry)XX  
DE Reagent of GTP-binding protein (10).XX  
KW G-protein; signal transducing protein; cell proliferation;  
KW cancer; dementia; receptor; transmission.XX  
OS Synthetic.XX  
PN JP02069665-A.XX  
PD 08-MAR-1990.XX  
PF 06-SEP-1988; 88JP-0221446.XX  
PR 06-SEP-1988; 88JP-0221446.XX  
PA (NISH/) NISHIMOTO I.XX  
DR WPI; 1990-119771/16.XX  
PT Reagent of GTP-binding protein - consists of polypeptide of less  
PT than 33 aminoacid residues where 4 are basic aminoacid residues,  
PT etc.XX  
PS Disclosure; Fig 1; 4pp; Japanese.

XX CC A reagent for activating GTP-binding protein (G-protein) consists of  
XX CC a polypeptide of less than 33 amino acids, of which more than 4 are  
XX CC basic, and more than 20% are hydrophobic. The reagent is useful for  
XX CC the purification, identification, functional examination, etc. of  
XX CC G-protein, a signal transducing protein. It will be possible to  
XX CC study the mechanism of cell proliferation, anticancer action,  
XX CC prevention of dementia, by activating/inhibiting the G-protein,  
XX CC using the reagent.  
XX CC The peptides given in R04048-59, R05848, R09313 and indicated in  
XX CC the Features Table bind selectively, activating the G-protein. The  
XX CC peptides have the same config. as the part of the receptor which binds  
XX CC to G-protein and is effective for the transmission of signal information  
XX CC into the cell. The reagent can be used in the form of powder or soln.  
XX CC Affinity purification can be simply carried out by using ion-exchange  
XX CC resin beads such as Sepharose. For example, the reagent is stuck onto  
XX CC the beads and homogenate is passed through a column packed with the  
XX CC beads to obtain a given G-binding protein.

XX  
SQ Sequence 20 AA;

Query Match 34.8%; Score 32; DB 11; Length 20;

Best Local Similarity 53.8%; Pred. No. 54;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LPVPGVLLKEFTV 14

| | | | |

Db 7 laarnvlvkdtv 19

Search completed: May 8, 2001, 15:07:08  
Job time: 224 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:08:20 ; Search time 68.24 Seconds  
(without alignments)  
18.127 Million cell updates/sec

Title: US-09-165-546A-10  
Perfect score: 92  
Sequence: 1 PLPVPGLLKFTVSGNI 18

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues  
Total number of hits satisfying chosen parameters: 4929

Minimum DB seq length: 0  
Maximum DB seq length: 25  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_67:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	39.1	18	2 S13974	chlorophyll a/b-bi
2	30	32.6	20	2 S29817	cytochrome p450 2C
3	29	31.5	12	2 PN0170	alcohol dehydrogen
4	28	30.4	21	2 S58431	phosphatidylinosit
5	26	28.3	17	2 S59481	hydroxyproline-ric
6	26	28.3	19	2 S74114	3-hydroxyacyl-CoA
7	26	28.3	22	2 A37043	Ig light chain, po
8	25	27.2	20	2 S57286	translation elonga
9	25	27.2	22	2 P00697	hemagglutinin (imp
10	24	26.1	13	2 A60856	inhibin alpha chal
11	24	26.1	15	2 B35389	urease (EC 3.5.1.5
12	24	26.1	15	2 A36315	recycling receptor
13	24	26.1	17	4 I76673	hypothetical COII/
14	24	26.1	24	2 A24417	interphotoreceptor
15	24	26.1	24	2 S70329	omega 1-40 secalin
16	24	26.1	24	2 S38729	probable malate ca
17	23	25.0	10	2 PN0165	triose-phosphate 1
18	23	25.0	12	2 PN0663	dystrophin-associa
19	23	25.0	15	2 I46909	voltage-dependent
20	23	25.0	15	2 A56963	acid phosphatase (
21	23	25.0	16	2 S69361	carbamoyl-phosphat
22	23	25.0	18	2 PC2280	prolylendopeptidas
23	23	25.0	18	2 A32917	protein phosphatas
24	23	25.0	19	2 A47689	flagellar sheath p
25	23	25.0	22	2 S63674	28 K secretory pro
26	23	25.0	23	2 T44418	carbamoyl-phosphat
27	23	25.0	24	2 S08185	proteasome alpha c
28	23	25.0	24	2 S40139	T-cell receptor J-
29	23	25.0	25	2 S00329	beta-N-acetylgluco

30	22.5	24.5	23	2	S37491	hypothetical prote
31	22	23.9	13	2	S09395	hypothetical prote
32	22	23.9	19	2	S69153	Neb-collostatin -
33	22	23.9	20	1	LFBSTT	tet leader peptide
34	22	23.9	20	1	LFBSTT	tet leader peptide
35	22	23.9	20	2	S23742	tet leader peptide
36	22	23.9	20	2	PC1151	equinotoxin 1C - s
37	22	23.9	20	2	B34016	tenebrosin B - sea
38	22	23.9	21	2	A42762	multicatalytic end
39	22	23.9	24	2	A33434	calcium-binding pr
40	22	23.9	24	2	S70333	endosperm protein,
41	22	23.9	25	2	S10850	alpha-amylase inhi
42	22	23.9	25	2	C25629	cytochrome-c oxida
43	22	23.9	25	2	A61457	alpha-glucosidase
44	21.5	23.4	12	2	PC4377	telomeric and tetr
45	21	22.8	10	2	A27617	triose-phosphate 1

ALIGNMENTS

RESULT 1  
S13974  
chlorophyll a/b-binding protein type I - garden pea (fragment)  
C:Species: Pisum sativum (garden pea)  
C:Date: 19-Mar-1997 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: S13974  
R:Jahns, P.; Junge, W.  
Eur. J. Biochem. 193, 731-736, 1990  
A:Title: Dicyclohexylcarbodiimide-binding proteins related to the short circuit of th  
A:Reference number: S13973; MUID:91065379  
A:Accession: S13974  
A:Molecule type: protein  
A:Residues: 1-18 <JAH>  
C:Genetics:  
A:Genome: nuclear  
C:Keywords: chlorophyll; chloroplast; light-harvesting complex; photosystem I; thylak

Query Match 39.1%; Score 36; DB 2; Length 18;  
Best Local Similarity 50.0%; Pred. No. 8.1;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 LPVPGVLLKFTVSGN 17  
Db 1 LAVPGILVPEALGLGN 16

RESULT 2

S29817  
cytochrome p450 2C23 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 05-Mar-1999  
C:Accession: S29817  
R:Marie, S.; Roussel, F.; Cresteil, T.  
Biochim. Biophys. Acta 1172, 124-130, 1993  
A:Title: Age- and tissue-dependent expression of CYP2C23 in the rat.  
A:Reference number: S29817; MUID:93176794  
A:Accession: S29817  
A:Molecule type: mRNA  
A:Residues: 1-20 <MAR>  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: heme; transmembrane protein

Query Match 32.6%; Score 30; DB 2; Length 20;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PLPVPGLLK 10  
Db 4 PLPIIGNLLE 13

```

RESULT 3
PN0170
alcohol dehydrogenase (EC 1.1.1.1) - fungus (Fusarium sporotrichioides) (fragment)
C:Species: Fusarium sporotrichioides
C:Date: 05-Aug-1994 #sequence_revision 06-Jan-1995 #text_change 31-Jan-1997
C:Accession: PN0170
R:Fukaya, N.; Chow, L.P.; Sugitara, Y.; Tsugita, A.; Ueno, Y.; Tabuchi, K.
submitted to JPID, May 1994
A:Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides
A:Reference number: PN0160
A:Accession: PN0170
A:Molecule type: protein
A:Residues: 1-12 <FUK>
A:Experimental source: strain M-1-1
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 31.5%; Score 29; DB 2; Length 12;
Best Local Similarity 66.7%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LLKFTVSG 16
   ||| :||
Db 2 LLKGYTDG 10

RESULT 4
S58431
phosphatidylinositol transfer protein isoform, 36K - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C:Accession: S58431
R:de Vries, K.J.; Heinrichs, A.A.J.; Cunningham, E.; Brunink, F.; Westerman, J.; Somerhaugen, J.
Biochem. J. 310, 643-649, 1995
A:Title: An isoform of the phosphatidylinositol-transfer protein transfers sphingomyelin
A:Reference number: S58430; MUID:95382786
A:Accession: S58431
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <DEV>
C:Superfamily: human phosphatidylinositol transfer protein

Query Match 30.4%; Score 28; DB 2; Length 21;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLLKEFTV 14
   ||||| :
Db 1 VLLKEYRV 8

RESULT 5
S59481
hydroxyproline-rich cell wall glycoprotein, 230K - kidney bean (fragment)
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 27-Apr-1996 #sequence_revision 19-Jul-1996 #text_change 05-Dec-1998
C:Accession: S59481
R:Wojtaszek, P.; Tretow, J.; Bolwell, G.P.
Plant Mol. Biol. 28, 1075-1087, 1995
A:Title: Specificity in the immobilisation of cell wall proteins in response to different
A:Reference number: S59481; MUID:96011753
A:Accession: S59481
A:Molecule type: protein
A:Residues: 1-17 <WOJ>
C:Keywords: glycoprotein; hydroxyproline
F:6,8,9,10,11/Modified site: hydroxyproline (Pro) #status experimental

Query Match 28.3%; Score 26; DB 2; Length 17;
Best Local Similarity 41.7%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 PLPVPGVLLKEF 12
   ||| :|
Db 6 PVPVPVVVYPTF 17

RESULT 6
S74114
3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C:Accession: S74114
R:Dieuaide-Noubhani, M.; Novikov, D.; Baumgart, E.; Vanhooren, J.C.T.; Fransen, M.; Gurr, J.
Eur. J. Biochem. 240, 660-666, 1996
A:Title: Further characterization of the peroxisomal 3-hydroxyacyl-CoA dehydrogenases
e acids di- and tri-hydroxycoprostanic acids are metabolized by separate multifunctional
A:Reference number: S74113; MUID:97008958
A:Accession: S74114
A:Molecule type: protein
A:Residues: 1-19 <DIE>
A:Experimental source: liver
C:Superfamily: 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase homol
C:Keywords: fatty acid beta-oxidation; mitochondrion; NAD; oxidoreductase

Query Match 28.3%; Score 26; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 VLLKEFTVSG 16
   ||| :||
Db 4 ILIKHVTIVG 13

RESULT 7
A37043
Ig light chain, polyclonal pool - elephantfish (Callorhynchus callorhynchus) (fragment)
C:Species: Callorhynchus callorhynchus
C:Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 16-Aug-1996
C:Accession: A37043
R:De Ioannes, A.E.; Aquila, H.L.
Immunogenetics 30, 175-180, 1989
A:Title: Amino terminal sequence of heavy and light chains from ratfish immunoglobulin
A:Reference number: A37043; MUID:89379282
A:Accession: A37043
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <DEI>
C:Keywords: immunoglobulin

Query Match 28.3%; Score 26; DB 2; Length 22;
Best Local Similarity 62.5%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PVPGVLLK 10
   |||| :
Db 12 PVPGNVTK 19

RESULT 8
S57286
translation elongation factor eEF-1 beta - Sulfolobus solfataricus (fragments)
C:Species: Sulfolobus solfataricus
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S57286
R:Arcari, P.; Raimo, G.; Ianniciello, G.; Gallo, M.; Bocchini, V.
Biochim. Biophys. Acta 1263, 86-88, 1995
A:Title: The first nucleotide sequence of an archaeal elongation factor 1-beta gene.
A:Reference number: S57268; MUID:95359209
A:Accession: S57286
A>Status: preliminary
A:Molecule type: protein

```

A:Residues: 1-20 <ARC>

Query Match 27.2%; Score 25; DB 2; Length 20;  
Best Local Similarity 60.0%; Pred. No. 7.3e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 VLLKEFTVSG 16  
|:| | | | |  
Db 5 VVLKVPVWG 14

RESULT 9

PQ0697  
hemagglutinin [imported] - rice (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: PQ0697  
R:Komatsu, S.; Kajiura, H.; Hirano, H.  
Theor. Appl. Genet. 86, 935-942, 1993  
A:Title: A rice protein library; a data-file of rice proteins separated by two-dimension  
A:Reference number: PQ0697  
A:Accession: PQ0697  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <KOM>

Query Match 27.2%; Score 25; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 8.1e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 EFTVSGNI 18  
|:| | | | |  
Db 3 KPVVGGNL 10

RESULT 10

A60856  
inhibin alpha chain - sheep (fragment)  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: A60856  
R:Leversha, L.J.; Robertson, D.M.; de Vos, F.L.; Morgan, F.J.; Hearn, M.T.W.; Wettenhall  
J. Endocrinol. 113, 213-221, 1987  
A:Title: Isolation of inhibin from ovine follicular fluid.  
A:Reference number: A60856; MUID:87224684  
A:Accession: A60856  
A:Molecule type: protein  
A:Residues: 1-13 <LEV>  
C:Comment: Inhibin suppresses follicle-stimulating hormone secretion.  
C:Superfamily: inhibin  
C:Keywords: disulfide bond; glycoprotein; gonad; heterodimer; hormone

Query Match 26.1%; Score 24; DB 2; Length 13;  
Best Local Similarity 80.0%; Pred. No. 6.7e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVP 5  
| | | |  
Db 4 PLXP 8

RESULT 11

B35389  
urease (EC 3.5.1.5) 15K chain - Morganella morganii (fragment)  
C:Species: Morganella morganii  
C:Date: 31-Aug-1990 #sequence\_revision 31-Aug-1990 #text\_change 23-Jun-1993  
C:Accession: B35389  
R:Hu, L.T.; Nicholson, E.B.; Jones, B.D.; Lynch, M.J.; Mobley, H.L.T.  
J. Bacteriol. 172, 3073-3080, 1990  
A:Title: Morganella morganii urease: purification, characterization, and isolation of ge

A:Reference number: A35389; MUID:90264298

A:Accession: B35389  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <HUA>  
C:Keywords: hydrolase

Query Match 26.1%; Score 24; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 7.8e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLPVPGLV 8  
| | | | |  
Db 6 PTPPLGGVI 13

RESULT 12

A36315  
recycling receptor pl80 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 30-Sep-1993  
C:Accession: A36315  
R:Isacke, C.M.; van der Geer, P.; Hunter, T.; Trowbridge, I.S.  
Mol. Cell. Biol. 10, 2606-2618, 1990  
A:Title: pl80, a novel recycling transmembrane glycoprotein with restricted cell type  
A:Reference number: A36315; MUID:90258846  
A:Accession: A36315  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <ISA>

Query Match 26.1%; Score 24; DB 2; Length 15;  
Best Local Similarity 62.5%; Pred. No. 7.8e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPVPGLV 9  
| | | | |  
Db 7 LPEPNVPL 14

RESULT 13

I76673  
hypothetical COII/ND5 mutant fusion protein - mouse mitochondrion (fragment)  
N:Alternate names: COII/ND5 protein  
C:Species: mitochondrion Mus musculus (house mouse)  
C:Date: 12-Aug-1996 #sequence\_revision 16-Jul-1998 #text\_change 20-Apr-2000  
C:Accession: I76673; I76674  
R:Nelson, I.; Gerasimov, S.; Marsac, C.; Lestienne, P.; Boursot, P.  
Mamm. Genome 4, 680-683, 1993  
A:Title: Sequence analysis of a deleted mitochondrial DNA molecule in heteroplasmic m  
A:Reference number: I57011; MUID:94108239  
A:Accession: I76673  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-17 <NEL1>  
A:Cross-references: GB:S68119; NID:g544777  
A:Accession: I76674  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 8-17 <NEL2>  
A:Cross-references: GB:S68119; NID:g544777  
C:Comment: This is the hypothetical translation of a sequence believed to result from  
C:Genetics:  
A:Gene: COII/ND5  
A:Genome: mitochondrion  
A:Genetic code: SGC1  
C:Keywords: fusion protein; mitochondrion  
F:1-7/Region: cytochrome-c oxidase chain II  
F:8-17/Region: NADH dehydrogenase (ubiquinone) chain 5

```
Query Match      26.1%; Score 24; DB 4; Length 17;
Best Local Similarity 50.0%; Pred. No. 9e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 LKFTVSGNI 18
   || | : ||
Db 5 LKPLTTNNI 14

RESULT 14
A24417
interphotoreceptor retinoid-binding protein - sheep (fragment)
N:Alternate names: Interstitial retinol-binding protein
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Nov-1997
C:Accession: A24417
R:Fong, S.L.; Cook, R.G.; Alvarez, R.A.; Liou, G.I.; Landers, R.A.; Bridges, C.D.B.
FEBS Lett. 205, 309-312, 1986
A:Title: N-terminal sequence homologies in interstitial retinol-binding proteins from 10
A:Reference number: A91365; MUID:86301171
A:Accession: A24417
A:Molecule type: protein
A:Residues: 1-24 <FON>
C:Superfamily: Interphotoreceptor retinoid-binding protein
C:Keywords: duplication

Query Match      26.1%; Score 24; DB 2; Length 24;
Best Local Similarity 35.3%; Pred. No. 1.3e+03;
Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPVPGVLLKFTVSGNI 18
   | : ||| : | |
Db 7 LDMAQVLLDNYTFPENL 23

RESULT 15
S70329
omega 1-40 secalin - rye (fragment)
C:Species: Secale cereale (rye)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 02-Jul-1998
C:Accession: S70329
R:Roche, A.; Calero, M.; Soriano, F.; Mendez, E.
Biochim. Biophys. Acta 1295, 13-22, 1996
A:Title: Identification of major rye secalins as coeliac immunoreactive proteins.
A:Reference number: S70327; MUID:96283789
A:Accession: S70329
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <ROC>
C:Superfamily: gliadin

Query Match      26.1%; Score 24; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PVPG 6
   ||||
Db 16 PVPG 19

Search completed: May 8, 2001, 15:08:21
Job time: 292 sec
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:07 ; Search time 40.07 Seconds  
(without alignments)  
15.388 Million cell updates/sec

Title: US-09-165-546A-10

Perfect score: 92

Sequence: 1 PLPVPVLLKEFTVSGNI 18

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34253486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	29.3	13	YC1A_SALTY	P25944 salmonella
2	25	27.2	15	UC29_MAIZE	P80635 zea mays (m
3	24	26.1	15	URE2_MORMO	P17338 morganella
4	24	26.1	24	DHAG_COMTE	P80705 comamonas t
5	24	26.1	24	IRBP_SHEEP	P12663 ovis aries
6	23	25.0	13	CRBL_VESMA	P17232 vespa manda
7	23	25.0	13	TEM_D_RANTE	P56919 rana tempor
8	23	25.0	20	COGA_CHIOP	P34156 chionoecete
9	22	23.9	13	TEMC_RANTE	P56918 rana tempor
10	22	23.9	13	TEME_RANTE	P56920 rana tempor
11	22	23.9	19	COOT_SARBU	Q09148 sarcophaga
12	22	23.9	20	KORA_METTM	P80904 methanobact
13	22	23.9	20	LPTR_BACST	P05658 bacillus st
14	22	23.9	20	LPTR_BACSU	P23053 bacillus su
15	22	23.9	20	TENB_ACTTE	P30834 actinia ten
16	22	23.9	25	COX6_NEUCR	Q01359 neurospora
17	21	22.8	10	TPIS_NICPL	P19118 nicotiana p
18	21	22.8	12	OP53_DROVI	P17645 drosophila p
19	21	22.8	13	RP37_LEUMA	P81754 leucophaea
20	21	22.8	13	CRBL_VESTR	P17231 vespa tropi
21	21	22.8	13	CRTC_BOVIN	P28489 bos taurus
22	21	22.8	19	FIBA_SHEEP	P14451 ovis aries
23	21	22.8	19	MIFH_TRISP	P81529 trichinella
24	21	22.8	20	KORC_METTM	P80906 methanobact
25	21	22.8	21	DCMS_PSECA	P19921 pseudomonas
26	21	22.8	22	VGLG_RABVA	P15199 rabies viru
27	21	22.8	24	CRTC_CANFA	P28490 canis fami
28	20.5	22.3	22	APCI_MACFA	P18657 macaca fasc
29	20	21.7	14	IF2G_RAT	P81795 rattus norv
30	20	21.7	16	ODPB_SOLTU	P81419 solanum tub
31	20	21.7	17	A45K_MYCBO	P80069 mycobacteri
32	20	21.7	17	TPIS_PINPS	P81666 pinus pinas
33	20	21.7	19	MIFH_TRIMR	P81530 trichuris m

34 20 21.7 23 1 GLMS\_CLOCO P80078 clostridium  
35 20 21.7 24 1 HEMT\_LINRE P23543 lingula ree  
36 20 21.7 24 1 LPER\_STRFR P45440 streptomyce  
37 20 21.7 25 1 LYC\_ASTRU P37715 asterias ru  
38 20 21.7 25 1 PATX\_PARC V P41736 paraponera  
39 19 20.7 9 1 MGMT\_BOVIN P29177 bos taurus  
40 19 20.7 10 1 PAPI\_PARMA P81863 pardachirus  
41 19 20.7 16 1 FOR2\_MYRGU P81437 myrmecia gu  
42 19 20.7 17 1 BOLA\_MEGPE P07495 megabombus  
43 19 20.7 17 1 NU4M\_TRIRU Q36834 trichophyto  
44 19 20.7 19 1 UP24\_UPEIN P82030 uperoleia 1  
45 19 20.7 20 1 AROQ\_ANYME P46380 amycolatops

#### ALIGNMENTS

RESULT 1  
YC1A\_SALTY  
ID YC1A\_SALTY STANDARD; PRT; 13 AA.  
AC P25944;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE PUTATIVE ACYL-COA THIOESTER HYDROLASE YC1A (EC 3.1.2.-) (P14 PROTEIN)  
DE (FRAGMENT).  
GN YC1A.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91094049; PubMed=2266561;  
RA Hannavy K., Barr G.C., Dorman C.J., Adamson J., Mazengera L.R.,  
RA Gallagher M.P., Evans J.S., Levine B.A., Trayer I.P., Higgins C.F.;  
RT "TonB protein of Salmonella typhimurium. A model for signal  
transduction between membranes.";  
RL J. Mol. Biol. 216:897-910(1990).  
CC - SIMILARITY: BELONGS TO THE ACYL COENZYME A HYDROLASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
CC EMBL; X56434; ; NOT\_ANNOTATED\_CDS.  
DR StyGene; SG10418; YC1A.  
KW Hypothetical protein; Hydrolase.  
FT NON\_TER 1  
SQ SEQUENCE 13 AA; 1307 MW; 30960DC5EFAC7417 CRC64;

Query Match 29.3%; Score 27; DB 1; Length 13;  
Best Local Similarity 83.3%; Pred. No. 11e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVPVG 6  
|||||  
Db 8 PLPVPVG 13

RESULT 2  
UC29\_MAIZE  
ID UC29\_MAIZE STANDARD; PRT; 15 AA.  
AC P80635;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 45)

DE (FRAGMENT).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
 OC Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Coleoptile;  
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
 RA Pernollet J.-C., Zivy M., de Vienne D.;  
 RT "The maize two dimensional gel protein database: towards an integrated  
 RT genome analysis program";  
 RL Theor. Appl. Genet. 93:997-1005(1996).  
 CC -|- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.9, ITS MW IS: 37.6 KDA.  
 DR Maize-2DPAGE; P80635; COLEOPTILE.  
 DT MaizedB; 123960; -. 1  
 FT NON\_TER 15 15  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1679 MW; 3D53086B16018BC1 CRC64;

Query Match 27.2%; Score 25; DB 1; Length 15;  
 Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPVP 5  
 Db 4 PVPIP 8  
 I: I: I  
 I: I: I

RESULT 3  
 URE2\_MORMO  
 ID URE2\_MORMO STANDARD; PRT; 15 AA.  
 AC P17338;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE UREASE BETA SUBUNIT (RC 3.5.1.5) (15 KDA SUBUNIT) (UREA  
 DE AMIDHYDROLASE) (FRAGMENT).  
 GN UREB.  
 OS Morganella morganii (Proteus morganii).  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Morganella.  
 OX NCBI\_TaxID=582;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90264298; PubMed=2345135;  
 RA Hu L.-T., Nicholson E.B., Jones B.D., Lynch M.J., Mobley H.L.T.;  
 RT "Morganella morganii ureases: purification, characterization, and  
 RT isolation of gene sequences.";  
 RL J. Bacteriol. 172:3073-3080(1990).  
 CC -|- CATALYTIC ACTIVITY: UREA + H(2)O = CO(2) + 2 NH(3).  
 CC -|- SUBUNIT: (ALPHA, BETA, GAMMA)(3) (BY SIMILARITY).  
 CC -|- SIMILARITY: TO OTHER UREASES BETA SUBUNITS.  
 DR PIR; B35389; B35389.  
 KW Hydrolase.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1530 MW; 2D98944F2F20C7E8 CRC64;

Query Match 26.1%; Score 24; DB 1; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLPVPGLV 8  
 Db 6 PTPLGVI 13  
 I: I: I  
 I: I: I

ID DHAG\_COMTE STANDARD; PRT; 24 AA.  
 AC P80705;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE ALDEHYDE DEHYDROGENASE, GAMMA CHAIN (EC 1.2.99.3) (ALDH) (FRAGMENT).  
 OS Comamonas testosteroni (Pseudomonas testosteroni).  
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.  
 OX NCBI\_TaxID=285;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=ATCC 15667;  
 RA Luykx D.M.A.M., Kim S.W., de Vries S., Duine J.A.;  
 RL Submitted (JUL-1996) to the SWISS-PROT data bank.  
 CC -|- CATALYTIC ACTIVITY: AN ALDEHYDE + ACCEPTOR + H(2)O = AN ACID +  
 CC REDUCED ACCEPTOR.  
 CC -|- COFACTOR: MOLYBDENUM.  
 CC -|- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA  
 CC CHAIN.  
 KW Oxidoreductase; Molybdenum.  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2585 MW; 9E66B518130EA938 CRC64;

Query Match 26.1%; Score 24; DB 1; Length 24;  
 Best Local Similarity 66.7%; Pred. No. 7.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 EFTVSG 16  
 Db 4 QFTVNG 9  
 I: I: I  
 I: I: I

RESULT 5  
 IRBP\_SHEEP  
 ID IRBP\_SHEEP STANDARD; PRT; 24 AA.  
 AC P12663;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE INTERPHOTORECEPTOR RETINOID-BINDING PROTEIN (IRBP) (INTERSTITIAL  
 DE RETINOL-BINDING PROTEIN) (FRAGMENT).  
 GN RBP3.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=86301171; PubMed=3743780;  
 RA Fong S.L., Cook R.G., Alvarez R.A., Liou G.I., Landers R.A.,  
 RA Bridges C.D.B.;  
 RT "N-terminal sequence homologies in interstitial retinol-binding  
 RT proteins from 10 vertebrate species.";  
 RL FEBS Lett. 205:309-312(1986).  
 CC -|- FUNCTION: IRBP SHUTTLES 11-CIS AND ALL TRANS RETINOIDS BETWEEN  
 CC THE RETINOL ISOMERASE IN THE PIGMENT EPITHELIUM AND THE VISUAL  
 CC PIGMENTS IN THE PHOTORECEPTOR CELLS OF THE RETINA.  
 CC -|- SUBCELLULAR LOCATION: INTERPHOTORECEPTOR MATRIX THAT PERMEATES  
 CC THE SPACE BETWEEN THE RETINA AND THE CONTIGUOUS LAYER OF PIGMENT  
 CC EPITHELIUM CELLS.  
 DR PIR: A24417; A24417.  
 KW Vitamin A; Transport.  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2799 MW; 02E6B61A8E4523 CRC64;

Query Match 26.1%; Score 24; DB 1; Length 24;  
 Best Local Similarity 35.3%; Pred. No. 7.2e+02;  
 Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPVFGVLLKFTVSGNI 18

Matches	5;	Conservative	1;	Mismatches	2;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

[illegible]

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1ZA (ZINC METALLOPROTEASES); ALSO KNOWN AS THE ASTACIN SUBFAMILY.  
 HSSP: P07584; 1IAD.  
 MEROPS: M12.001; --.  
 InterPro: IPR000130; --.  
 PROSITE: PS00142; ZINC\_PROTEASE; PARTIAL.  
 Hydrolase; Metalloprotease; Zinc; Collagen degradation.  
 NON\_TER 20  
 SEQUENCE 20 AA; 2108 MW; 2BC7A93D02A97D8 CRC64;

```

Query Match          25.0%; Score 23; DB 1; Length 20;
Best Local Similarity 33.3%; Pred. No. 8.8e+02;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

y 7 VLLKEFTVSGNI 18
   :| 1: 11:
b 3 ILQDEYLVSGGV 14

RESULT 9
EMC_RANTE EMC_C_RANTE
D TEMC_RANTE STANDARD; PRT; 13 AA.
C P56918:
T T 30-MAY-2000 (Rel. 39, Created)
T 01-OCT-2000 (Rel. 40, Last sequence update)
T 01-OCT-2000 (Rel. 40, Last annotation update)

```

temporaria.;  
Eur. J. Biochem. 242:788-792(1996).  
-!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM- POSITIVE

```

CC BACTERIA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
CC Amphibian skin; Antibiotic; Amidation; Multigene family.
KW MOD_RES 13 13 AMIDATION.
FT SEQUENCE 13 AA; 1363 MW; 2201403A655B2448 CRC64;
SQ

Query Match 23.9%; Score 22; DB 1; Length 13;
Best Local Similarity 62.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPVPGVLL 9
Db 2 LPILGNLL 9

RESULT 10
TEMP_RANTE STANDARD; PRT; 13 AA.
AC P56920;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TEMPORIN E.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=97175050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Barra D.,
RA "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria."
RL Eur. J. Biochem. 242:788-792(1996).
CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE
CC BACTERIA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
CC Amphibian skin; Antibiotic; Amidation; Multigene family.
KW MOD_RES 13 13 AMIDATION.
FT SEQUENCE 13 AA; 1379 MW; 26505DFA79A92448 CRC64;
SQ

Query Match 23.9%; Score 22; DB 1; Length 13;
Best Local Similarity 62.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPVPGVLL 9
Db 2 LPILGNLL 9

RESULT 11
COOT_SARBU STANDARD; PRT; 19 AA.
AC Q09148;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NEB-COLLOSTATIN (FOLLICULOSTATIN).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidae; Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7385;
RN [1]
RP SEQUENCE.

```

```

RX MEDLINE=95188911; PubMed=7883009;
RA Bylemans D., Proost P., Samijn B., Borovsky D., Grauwels L.,
RA Huybrechts R., van Damme J., van Beeumen J., de Loof A.;
RT "Neb-collostatin, a second folliculostatins of the grey fleshfly,
RT Neobellieria bullata."
RL Eur. J. Biochem. 228:45-49(1995).
CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. IT INHIBITS YOLK SYNTHESIS SO
CC INHIBITING THE SUBSEQUENT YOLK DEPOSITION IN PREVITELLOGENIC
CC FOLLICLES.
CC -!- SIMILARITY: TO THE NONHELICAL REGIONS OF THE COLLAGEN FAMILY.
CC -!- CAUTION: NEB-COLLOSTATIN MAY BE PROTEOLYTICALLY CLEAVED FROM
CC COLLAGEN IV.
SQ SEQUENCE 19 AA; 1880 MW; 41B6AF5F0CEB8251 CRC64;

Query Match 23.9%; Score 22; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPVP 5
Db 7 LPVP 10

RESULT 12
KORA_METTM STANDARD; PRT; 20 AA.
AC P80904;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2-OXOGLUTARATE SYNTHASE SUBUNIT KORA (EC 1.2.7.3) (2-KETOGLUTARATE
DE OXIDOREDUCTASE ALPHA CHAIN) (KOR) (2-OXOGLUTARATE-FERREDOXIN
DE OXIDOREDUCTASE ALPHA SUBUNIT) (FRAGMENT).
GN KORA.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE.
RX MEDLINE=97261844; PubMed=9108258;
RA "Streegen A., Linder D., Thauer R.K., Hedderich R.;
RT "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT in Methanobacterium thermoautotrophicum."
RL Eur. J. Biochem. 244:862-868(1997).
CC -!- CATALYTIC ACTIVITY: 2-OXOBUTANOATE + COA + OXIDIZED FERREDOXIN =
CC PROPANOYL-COA + CO(2) + REDUCED FERREDOXIN.
CC -!- SUBUNIT: HETEROTETRAMER OF THE KORA, KORB, KORC AND KORD
CC SUBUNITS.
CC -!- MISCELLANEOUS: THE PH OPTIMUM IS PH 9.0 AND THE OPTIMAL
CC TEMPERATURE IS 70 DEGREES CELSIUS.
KW Oxidoreductase.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2080 MW; 112E7E0E4AC76B76 CRC64;

Query Match 23.9%; Score 22; DB 1; Length 20;
Best Local Similarity 37.5%; Pred. No. 1.3e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 KEFTVSGN 17
Db 2 EEYFIQGN 9

RESULT 13
LPTR_BACST STANDARD; PRT; 20 AA.
AC P05658;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)

```

DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.  
GN TETL  
OS Bacillus stearothermophilus, Bacillus cereus, and  
OS Staphylococcus hyicus.  
OG Plasmid pTH15, Plasmid pBC16, and Plasmid pSTEL.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1422, 1396, 1284;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PLASMID=pTH15;  
RX MEDLINE=86031344; PubMed=2996983;  
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;  
RT "Nucleotide sequence of the tetracycline resistance gene of pTH15, a  
RT thermophilic Bacillus plasmid: comparison with staphylococcal Tcr  
RT controls.";  
RL Gene 37:131-138(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC PLASMID=pBC16;  
RX MEDLINE=90221899; PubMed=2109312;  
RA Palva A., Vidgren G., Simonen M., Rintala H., Laamanen P.;  
RT "Nucleotide sequence of the tetracycline resistance gene of pBC16  
RT from Bacillus cereus.";  
RL Nucleic Acids Res. 18:1635-1635(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.hyicus; PLASMID=pSTEL;  
RX MEDLINE=92321725; PubMed=1622166;  
RA Schwarz S., Cardoso M., Wegener H.C.;  
RT "Nucleotide sequence and phylogeny of the tet(L) tetracycline  
RT resistance determinant encoded by plasmid pSTEL from Staphylococcus  
RT hyicus.";  
RL Antimicrob. Agents Chemother. 36:580-588(1992).  
RN [4]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; D00006; BAA00004.1; -;  
DR EMBL; M1036; AAA22850.1; -;  
DR EMBL; X51366; CAA35750.1; -;  
DR EMBL; X60828; CAA43219.1; -;  
DR PIR; S09233; LFBSTU.  
DR PIR; S23742; S23742.  
KW Leader peptide; Antibiotic resistance; Plasmid.  
SQ SEQUENCE 20 AA; 2253 MW; 18D0F4A6CA231CA1 CRC64;

Query Match 23.9%; Score 22; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 1.3e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 VLLKEFTVS 15  
| | | | : | |  
Db 9 VQLKEGSVS 17

RESULT 14  
ID LPTR\_BACSU STANDARD; PRT; 20 AA.  
AC P23053;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.  
GN TETL  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=B9000797; PubMed=2844262;  
RA Sakaguchi R., Amano H., Shishido K.;  
RT "Nucleotide sequence homology of the tetracycline-resistance  
RT determinant naturally maintained in Bacillus subtilis Marburg 168  
RT chromosome and the tetracycline-resistance gene of B. subtilis  
RT plasmid pNS1981.";  
RL Biochim. Biophys. Acta 950:441-444(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91257555; PubMed=1646143;  
RA Amano H., Sakaguchi R., Shishido K.;  
RT "An insertion of Escherichia coli transposable element ISLK into the  
RT site immediately before tetracycline-resistance determinant of  
RT Bacillus subtilis chromosomal DNA fragment in cloning in E. coli.";  
RL FEMS Microbiol. Lett. 63:5-8(1991).  
RN [3]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; X08034; CAA30826.1; -;  
DR EMBL; X58999; CAA41744.1; -;  
DR EMBL; Z99124; CABI6115.1; -;  
DR PIR; S04802; LFBSTT.  
DR PIR; S17282; S17282.  
DR Subtilist; BG11051; tetL.  
KW Leader peptide; Antibiotic resistance.  
SQ SEQUENCE 20 AA; 2298 MW; E870F516CA23111A CRC64;

Query Match 23.9%; Score 22; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 1.3e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 VLLKEFTVS 15  
| | | | : | |  
Db 9 VQLKEGSVS 17

RESULT 15  
ID TENB\_ACTTE STANDARD; PRT; 20 AA.  
AC P30834;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE TENEBROSIN B (FRAGMENT).  
OS Actinia tenebrosa (Australian red waratah sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
OC Nynanthaeae; Actinellidae; Actinia.  
OX NCBI\_TaxID=6105;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=90232538; PubMed=1970442;  
RA Norton R.S., Bobek G., Ivanov J.O., Thomson M., Fiala-Beer E.,  
RA Moritz R.L., Simpson R.J.;  
RT "Purification and characterisation of proteins with cardiac  
RT stimulatory and haemolytic activity from the anemone Actinia  
RT tenebrosa.";  
RL Toxicon 28:29-41(1990).  
CC -I- FUNCTION: THIS CARDIAC STIMULATORY AND HEMOLYTIC PROTEIN IS A  
CC -I- CHANNEL-FORMING AND/OR MEMBRANE-PENETRATING PROTEIN.  
CC -I- SIMILARITY: BELONGS TO THE TENEBROSIN FAMILY.  
DR PIR; B34016; B34016.

KW Hemolysis; Toxin; Transmembrane.

FT NON\_TER 20 20

SQ SEQUENCE 20 AA: 1960 MW: FA32B426009FF5FA CRC64;

Query Match 23.9%; Score 22; DB 1; Length 20;

Best Local Similarity 26.7%; Pred. No. 1.3e+03;

Matches 4; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 4 VPGVLLKEFTVSGNI 18

| | : : | : | :

Db 4 VAGAVIEGATLTENV 18

Search completed: May 8, 2001, 15:14:08

Job time: 529 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:16:06 ; Search time 114.89 Seconds  
(without alignments)  
18.363 Million cell updates/sec

Title: US-09-165-546A-10  
Perfect score: 92  
Sequence: 1 PLPVPGLLKFTVSGNI 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_15:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_Organelle:\*  
9: sp\_Phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	32.6	20	11	Q9QW31
2	29	31.5	21	6	Q9TR36
3	29	31.5	22	2	Q9ZAW2
4	29	31.5	24	2	O52390
5	28	30.4	21	6	Q9TR37
6	28	30.4	24	4	Q9UCP1
7	28	30.4	24	11	Q9QVC8
8	27.5	29.9	20	6	Q9TRU5
9	27.5	29.9	24	2	Q9R558
10	27	29.3	18	6	P79214
11	25	27.2	15	4	Q9UC22
12	25	27.2	16	2	Q47605
13	25	27.2	17	4	Q9UC89
14	25	27.2	21	9	Q9T167
15	25	27.2	25	10	O49883
16	24	26.1	20	2	Q9R502
17	24	26.1	21	3	Q9UR89
18	24	26.1	21	11	Q9Z167
19	24	26.1	23	11	Q9Z208

20	24	26.1	24	4	Q16475	Q16475 homo sapien
21	24	26.1	24	4	Q9UMB2	Q9UMB2 homo sapien
22	24	26.1	24	5	P92152	P92152 caenorhabdi
23	24	26.1	24	10	Q9S8A6	Q9S8A6 secale cere
24	24	26.1	25	14	Q9YN05	Q9YN05 myxoma viru
25	23	25.0	9	4	Q9P225	Q9P225 homo sapien
26	23	25.0	11	11	Q9QYF6	Q9QYF6 mus musculu
27	23	25.0	15	2	O69173	O69173 yersinia pe
28	23	25.0	15	6	Q28822	Q28822 oryctolagus
29	23	25.0	17	6	Q9TRU8	Q9TRU8 bos taurus
30	23	25.0	17	14	Q9IHJ0	Q9IHJ0 human polio
31	23	25.0	17	14	Q9IHJ9	Q9IHJ9 human polio
32	23	25.0	17	14	Q9IH18	Q9IH18 human polio
33	23	25.0	17	14	Q9IH17	Q9IH17 human polio
34	23	25.0	17	14	Q9IH16	Q9IH16 human polio
35	23	25.0	17	14	Q9IH15	Q9IH15 human polio
36	23	25.0	17	14	Q9IH14	Q9IH14 human polio
37	23	25.0	17	14	Q9IH13	Q9IH13 human polio
38	23	25.0	17	14	Q9IH12	Q9IH12 human polio
39	23	25.0	17	14	Q9IH11	Q9IH11 human polio
40	23	25.0	17	14	Q9IH10	Q9IH10 human polio
41	23	25.0	17	14	Q9IHH9	Q9IHH9 human polio
42	23	25.0	17	14	Q9IHH8	Q9IHH8 human polio
43	23	25.0	17	14	Q9IHH7	Q9IHH7 human polio
44	23	25.0	17	14	Q9IHH6	Q9IHH6 human polio
45	23	25.0	17	14	Q9IHH5	Q9IHH5 human polio

## ALIGNMENTS

RESULT 1

Q9QW31 ID Q9QW31 PRELIMINARY; PRT; 20 AA.  
AC Q9QW31:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE CYP2C23-CYTCHROME P-450.  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93176794; PubMed=7679925;  
RA Marie S., Roussel F., Cresteil T.;  
RT "Age- and tissue-dependent expression of CYP2C23 in the rat.";  
RL Biochim. Biophys. Acta 1172:124-130(1993).  
SQ SEQUENCE 20 AA; 2128 MW; 490517DA2A032F65 CRC64;

Query Match 32.6%; Score 30; DB 11; Length 20;  
Best Local Similarity 60.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PLPVPGLLK 10

Db 4 PLPIGNLLE 13

RESULT 2

Q9TR36 ID Q9TR36 PRELIMINARY; PRT; 21 AA.  
AC Q9TR36:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE PHOSPHATIDYLINOSITOL-TRANSFER PROTEIN 36 KDA ISOFORM (FRAGMENT).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=95382786; PubMed=7654206;  
 RX de Vries K.J., Heinrichs A.A., Cunningham E., Brunink F.,  
 RT Westernman J., Somerharju P.J., Cockcroft S., Wirtz K.W., Snoek G.T.;  
 RA "An isoform of the phosphatidylinositol-transfer protein transfers  
 RT sphingomyelin and is associated with the Golgi system.";  
 RL Biochem. J. 310:643-649(1995).  
 DR INTERPRO: IPR001666; -;  
 DR PFAM: PF02121; IP-trans; 1.  
 SQ SEQUENCE 21 AA; 2442 MW; 0978BC2B0067EF0E CRC64;

Query Match 31.5%; Score 29; DB 6; Length 21;  
 Best Local Similarity 75.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLLKEFTV 14  
 Db 1 VLIKEPRV 8  
 |||||

RESULT 3  
 O92AW2 PRELIMINARY; PRT; 22 AA.  
 ID Q92AW2  
 AC Q92AW2  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE ENVA (FRAGMENT).  
 GN ENVA.  
 OS Francisella tularensis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Francisella group;  
 OC Francisella.  
 OX NCBI\_TaxID=263;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LVS (LIVE VACCINE STRAIN);  
 RA Salimnia H., Clairoux N., Bolssinot M.;  
 RT "Cloning and characterization of Francisella tularensis LVS homolog of  
 RT the E. coli cell division gene ftsZ.";  
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U76309; AAC99559.1; -;  
 FT NON\_TER 22  
 SQ SEQUENCE 22 AA; 2333 MW; ADA9026DE208C0A5 CRC64;

Query Match 31.5%; Score 29; DB 2; Length 22;  
 Best Local Similarity 55.8%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 LLKEFTVSG 16  
 Db 6 IAKFSVTG 14  
 :|||:|

RESULT 4  
 O52390 PRELIMINARY; PRT; 24 AA.  
 ID O52390  
 AC O52390  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE SERINE/THREONINE PHOSPHATASE 2B (FRAGMENT).  
 GN PP2B-CYANOL.  
 OS Nodularia spumigena.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nodularia.  
 OX NCBI\_TaxID=70799;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=L575;  
 RX MEDLINE=98047369; PubMed=9385145;

RA Shi L., Carmichael W.W.;  
 RT "pp1-cyanol, a protein serine/threonine phosphatase 1 gene from the  
 RL cyanobacterium Microcystis aeruginosa UTEX 2063.";  
 DR Arch. Microbiol. 168:528-531(1997).  
 DR EMBL: AF037161; AAC46049.1; -;  
 DR INTERPRO: IPR000934; -;  
 DR PFAM: PF00149; STphosphatase; 1.  
 FT NON\_TER 1  
 FT NON\_TER 24  
 SQ SEQUENCE 24 AA; 2768 MW; DB9C36D51C0DA51D CRC64;

Query Match 31.5%; Score 29; DB 2; Length 24;  
 Best Local Similarity 60.0%; Pred. No. 4.2e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 LLKEFTVSGN 17  
 Db 5 LKLFVGVGN 14  
 :|||:|

RESULT 5  
 Q9TR37 PRELIMINARY; PRT; 21 AA.  
 ID Q9TR37  
 AC Q9TR37  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE PHOSPHATIDYLINOSITOL-TRANSFER PROTEIN 35 KDA ISOFORM (FRAGMENT).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=95382786; PubMed=7654206;  
 RA de Vries K.J., Heinrichs A.A., Cunningham E., Brunink F.,  
 RA Westernman J., Somerharju P.J., Cockcroft S., Wirtz K.W., Snoek G.T.;  
 RT "An isoform of the phosphatidylinositol-transfer protein transfers  
 RT sphingomyelin and is associated with the Golgi system.";  
 RL Biochem. J. 310:643-649(1995).  
 DR INTERPRO: IPR001666; -;  
 DR PFAM: PF02121; IP-trans; 1.  
 SQ SEQUENCE 21 AA; 2462 MW; E5A50C9DA8C4F2D6 CRC64;

Query Match 30.4%; Score 28; DB 6; Length 21;  
 Best Local Similarity 75.0%; Pred. No. 5.3e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLLKEFTV 14  
 Db 1 VLLKEFRV 8  
 |||||

RESULT 6  
 Q9UCP1 PRELIMINARY; PRT; 24 AA.  
 ID Q9UCP1  
 AC Q9UCP1  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE FKBP59-59 KDA IMMUNOPHILIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=92285692; PubMed=1376003;  
 RA Tai P.K., Albers M.W., Chang H., Faber L.E., Schreiber S.L.;  
 RT "Association of a 59-kilodalton immunophilin with the glucocorticoid

RT receptor complex."  
RL Science 256:1315-1318(1992).  
SQ SEQUENCE 24 AA; 2486 MW; 7BCF45D0AF6EF736 CRC64;

Query Match 30.4%; Score 28; DB 4; Length 24;  
Best Local Similarity 71.4%; Pred. No. 6.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVPGV 7  
|||: ||  
Db 16 PLPMEGV 22

RESULT 7  
Q9QVC8 PRELIMINARY; PRT; 24 AA.  
AC Q9QVC8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE FKBP59-59 KDA IMMUNOPHTLIN.  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92285692; PubMed=1376003;  
RA Tai P.K., Albers M.W., Chang H., Faber L.E., Schreiber S.L.;  
RT "Association of a 59-kilodalton immunophilin with the glucocorticoid  
RT receptor complex."  
RL Science 256:1315-1318(1992).  
SQ SEQUENCE 24 AA; 2478 MW; 901A4B67B2AF5D16 CRC64;

Query Match 30.4%; Score 28; DB 11; Length 24;  
Best Local Similarity 71.4%; Pred. No. 6.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVPGV 7  
|||: ||  
Db 16 PLPLOGV 22

RESULT 8  
Q9TRU5 PRELIMINARY; PRT; 20 AA.  
AC Q9TRU5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE FACTOR IX LIGHT CHAIN (GLA DOMAIN).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92132840; PubMed=1776139;  
RA Rao L.V., Hoang A.D.;  
RT "Purification and characterization of rabbit factor IX and its  
RT existence as a two-chain factor IX alpha in circulating plasma."  
RL Thromb. Res. 64:57-68(1991).  
DR HSP: P00740; ICFH.  
SQ SEQUENCE 20 AA; 2317 MW; 9368E94B44BF5800 CRC64;

Query Match 29.9%; Score 27.5; DB 6; Length 20;  
Best Local Similarity 70.0%; Pred. No. 6.1e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 9 LKEFTVSGNI 18

Db |||: |||:  
6 LEEF-VSGNL 14

RESULT 9  
Q9R558 PRELIMINARY; PRT; 24 AA.  
AC Q9R558;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE SUPERANTIGEN (FRAGMENT).  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93352815; PubMed=8349810;  
RA Mollick J.A., Miller G.G., Musser J.M., Cook R.G., Grossman D.,  
RA Rich R.R.;  
RT "A novel superantigen isolated from pathogenic strains of  
RT Streptococcus pyogenes with aminoterminal homology to staphylococcal  
RT enterotoxins B and C".  
RL J. Clin. Invest. 92:710-719(1993).  
DR HSP: P34071; ISTE.  
SQ SEQUENCE 24 AA; 2563 MW; 7AE3BC2BA3A180B6 CRC64;

Query Match 29.9%; Score 27.5; DB 2; Length 24;  
Best Local Similarity 47.6%; Pred. No. 7.5e+02;  
Matches 10; Conservative 2; Mismatches 6; Indels 3; Gaps 2;

QY 1 PLPVPGVLLK--EFT-VSGNI 18  
|||: |||:  
Db 4 PDPTEQLNKSSQFTGVGMNL 24

RESULT 10  
P79214 PRELIMINARY; PRT; 18 AA.  
AC P79214;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
DE CFTR PROTEIN.  
GN CFTR.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Vuillaume S., Denamur E.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: X95931; CAA65172.1; -.  
SQ SEQUENCE 18 AA; 2080 MW; 4DIC55056C1CE1B5 CRC64;

Query Match 29.3%; Score 27; DB 6; Length 18;  
Best Local Similarity 58.3%; Pred. No. 6.6e+02;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 PLPVPGVLLKEF 12  
|||: |||:  
Db 5 PLEKAGVLSKLF 16

RESULT 11  
Q9UCC2 PRELIMINARY; PRT; 15 AA.  
ID Q9UCC2  
AC Q9UCC2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 35 KDA HEPARIN-RELEASEABLE PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=94059921; PubMed=8241100;
RA Novotny W.F., Maffi T., Mehta R.L., Milner P.G.;
RT "Identification of novel heparin-releasable proteins, as well as the
RT cytokines midkine and pleiotrophin, in human desheparin plasma.";
RL Arterioscler. Thromb. 13:1798-1805(1993).
SQ SEQUENCE 15 AA; 1454 MW; 0585FEF4FE8F6265 CRC64;

Query Match 27.2%; Score 25; DB 4; Length 15;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPVP 5
   | : | : |
Db 5 PVPIP 9

RESULT 12
Q47605 PRELIMINARY; PRT; 16 AA.
AC Q47605;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE C (FRAGMENT).
GN C.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139577; PubMed=1995588;
RA Tao T., Bourne J.C., Blumenthal R.M.;
RT "A family of regulatory genes associated with type II restriction-
RT modification systems.";
RL J. Bacteriol. 173:1367-1375(1991).
DR EMBL; M63622; AAA24561.1; -
FT NON_TER 1
SQ SEQUENCE 16 AA; 1853 MW; E46774511496607C CRC64;

Query Match 27.2%; Score 25; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 LKEFTVSGN 17
   | : | : |
Db 3 LLDFTTIGN 11

RESULT 13
Q9UC89 PRELIMINARY; PRT; 17 AA.
AC Q9UC89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ONCOFETAL-LAMININ BINDING [COLLAGEN ALPHA 1(III)CHAIN, OF-LB-COLLAGEN
DE ALPHA 1(III)].
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE.
RX MEDLINE=95169134; PubMed=7864881;
RA Minafra I.P., Andriolo M., Basirico L., Aquino A., Minafra S.,
RA Boutillon M.M., van der Rest M.;
RT "Onco-fetal/laminin-binding collagen from colon carcinoma: detection
RT of new sequences.";
RL Biochem. Biophys. Res. Commun. 207:852-859(1995).
SQ SEQUENCE 17 AA; 1566 MW; 8436525120A9B421 CRC64;

Query Match 27.2%; Score 25; DB 4; Length 17;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLPVP 6
   | : | : |
Db 6 PLGIPG 11

RESULT 14
Q9TI67 PRELIMINARY; PRT; 21 AA.
AC Q9TI67;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GP53.
OS Bacteriophage A118.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
OX NCBI_TaxID=40521;
RN [1]
RP SEQUENCE FROM N.A.
RA Loessner M.J.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96020653; PubMed=8577256;
RA Loessner M.J., Wendlinger G., Scherer S.;
RT "Heterogeneous endolysins in Listeria monocytogenes bacteriophages: a
RT new class of enzymes and evidence for conserved holin genes within the
RT siphoviral lysis cassettes.";
RL Mol. Microbiol. 16:1231-1241(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Loessner M.J., Inman R.B., Lauer P., Calendar R.;
RT "Complete nucleotide sequence, molecular analysis and genome structure
RT of bacteriophage A118 of Listeria monocytogenes: implications for
RT phage evolution.";
RL Mol. Microbiol. 35:324-340(2000).
DR EMBL; AJ242593; CAB53843.1; -
DR INTERPRO: IPR000217; -
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
SQ SEQUENCE 21 AA; 2485 MW; 7B5D940B/5D7CF90 CRC64;

Query Match 27.2%; Score 25; DB 9; Length 21;
Best Local Similarity 40.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 LKEFTVSGNI 18
   | : | : |
Db 1 MREIETIGNI 10

RESULT 15
O49883 PRELIMINARY; PRT; 25 AA.
AC O49883;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CF-4 RESISTANCE GENE CLUSTER (FRAGMENT).

```

GN LOXC HOMOLOGUE.  
OS Lycopersicon hirsutum.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
OC Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=62890;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. CF4;  
RX MEDLINE=98074802; PubMed=9413991;  
RA Parniske M., Hammond-Kosack K.E., Golstein C., Thomas C.M.,  
RA Jones D.A., Harrison K., Wulff B.B., Jones J.D.;  
RT "Novel disease resistance specificities result from sequence exchange  
RT between tandemly repeated genes at the Cf-4/9 locus of tomato."  
RL Cell 91:821-832(1997).  
DR EMBL; AJ002235; CAA05271.1; -.  
DR HSSP; P08170; 1YGE.  
FT NON\_TER 1  
SQ SEQUENCE 25 AA; 2710 MW; 72CDC71677C3792F CRC64;

Query Match 27.2%; Score 25; DB 10; Length 25;  
Best Local Similarity 46.2%; Pred. NO. 2.1e+03;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 4 VPGVLLKEFTVSG 16  
: | | | | | | | | | |  
Db 1 MPYELLKPSEPG 13

Search completed: May 8, 2001, 15:16:06  
Job time: 532 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:05:10 ; Search time 62.11 Seconds  
(without alignments)  
3.093 Million cell updates/sec

Title: US-09-165-546A-7

Perfect score: 57

Sequence: 1 LLMWITQCFL 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 110741

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2.6/ptodata/2/1aa/5A-COMB.pep.\*  
2: /cgn2.6/ptodata/2/1aa/5B-COMB.pep.\*  
3: /cgn2.6/ptodata/2/1aa/6A-COMB.pep.\*  
4: /cgn2.6/ptodata/2/1aa/6B-COMB.pep.\*  
5: /cgn2.6/ptodata/2/1aa/PCTUS-COMB.pep.\*  
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	11	4	US-09-183-931-42
2	47	82.5	9	4	US-09-183-931-43
3	35.5	62.3	20	2	US-08-934-915-84
4	33	57.9	9	4	US-09-183-931-44
5	31	54.4	14	2	US-08-764-640-35
6	31	54.4	14	2	US-08-764-640-36
7	31	54.4	14	3	US-08-973-225-35
8	31	54.4	14	3	US-08-973-225-36
9	31	54.4	14	4	US-09-244-298A-35
10	31	54.4	14	4	US-09-244-298A-36
11	28	49.1	16	5	PCT-US91-02942-14
12	27.5	48.2	24	2	US-08-470-419-2
13	27.5	48.2	24	2	US-08-761-828-2
14	27.5	48.2	24	2	US-08-290-105-2
15	27.5	48.2	24	3	US-08-776-949-2
16	27.5	48.2	24	3	US-08-482-810-2
17	27	47.4	19	5	PCT-US93-08435-61
18	27	47.4	20	2	US-08-483-636-6
19	27	47.4	20	2	US-08-483-632-6
20	27	47.4	20	5	PCT-US94-07659-10
21	26	45.6	6	2	US-08-482-228-186
22	26	45.6	6	3	US-08-482-528-186
23	26	45.6	11	1	US-08-451-947-45
24	26	45.6	11	2	US-08-424-826A-45
25	26	45.6	11	3	US-08-928-694-45
26	26	45.6	11	4	US-09-224-785-7
27	26	45.6	11	5	PCT-US91-06950-45

28	26	45.6	12	1	US-08-451-947-24	Sequence 24, Appl
29	26	45.6	12	2	US-08-424-826A-24	Sequence 24, Appl
30	26	45.6	12	3	US-08-928-694-24	Sequence 24, Appl
31	26	45.6	12	5	PCT-US91-06950-24	Sequence 24, Appl
32	26	45.6	14	1	US-08-311-611A-81	Sequence 81, Appl
33	26	45.6	14	1	US-08-372-783-81	Sequence 81, Appl
34	26	45.6	14	1	US-08-372-105-81	Sequence 81, Appl
35	26	45.6	14	1	US-08-306-473A-81	Sequence 81, Appl
36	26	45.6	14	1	US-08-209-762-81	Sequence 81, Appl
37	26	45.6	14	1	US-08-473-344-81	Sequence 81, Appl
38	26	45.6	14	2	US-08-803-79	Sequence 79, Appl
39	26	45.6	14	2	US-08-485-445A-81	Sequence 81, Appl
40	26	45.6	14	2	US-08-621-259A-25	Sequence 25, Appl
41	26	45.6	14	3	US-09-119-263-81	Sequence 81, Appl
42	26	45.6	14	4	US-08-657-162-81	Sequence 81, Appl
43	26	45.6	14	4	US-09-224-480-81	Sequence 81, Appl
44	26	45.6	14	5	PCT-US94-02465-81	Sequence 81, Appl
45	26	45.6	14	5	PCT-US95-00498-81	Sequence 81, Appl

#### ALIGNMENTS

RESULT 1  
US-09-183-931-42  
; Sequence 42, Application US/09183931C  
; Patent No. 6210886  
; GENERAL INFORMATION:  
; APPLICANT: Van Baren, Nicolas  
; APPLICANT: Brasseur, Francis  
; APPLICANT: Boon-Falleur, Thierry  
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING  
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS  
; FILE REFERENCE: LUD 5527.1-JEL/ES  
; CURRENT APPLICATION NUMBER: US/09/183-931C  
; CURRENT FILING DATE: 2000-02-28  
; EARLIER APPLICATION NUMBER: US 09/078,422  
; EARLIER FILING DATE: 1998 - 02 - 02  
; NUMBER OF SEQ ID NOS: 44  
; SEQ ID NO 42  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION:  
US-09-183-931-42

Query Match 100.0%; Score 57; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00087;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMWITQCFL 10  
| | | | | | | | | |  
DB 2 LLMWITQCFL 11

RESULT 2  
US-09-183-931-43  
; Sequence 43, Application US/09183931C  
; Patent No. 6210886  
; GENERAL INFORMATION:  
; APPLICANT: Van Baren, Nicolas  
; APPLICANT: Brasseur, Francis  
; APPLICANT: Boon-Falleur, Thierry  
; TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING  
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS  
; FILE REFERENCE: LUD 5527.1-JEL/ES  
; CURRENT APPLICATION NUMBER: US/09/183,931C  
; CURRENT FILING DATE: 2000-02-28  
; EARLIER APPLICATION NUMBER: US 09/018,422  
; EARLIER FILING DATE: 1998 - 02 - 04  
; NUMBER OF SEQ ID NOS: 44

SEQ ID NO 43  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION:  
US-09-183-931-43

Query Match 82.5%; Score 47; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLMWITOC 8  
Db 2 LLMWITOC 9

RESULT 3  
US-08-934-915-84  
Sequence 84, Application US/08934915  
Patent No. 5932412  
GENERAL INFORMATION:  
APPLICANT: DILLNER, JOAKIM  
APPLICANT: DILLNER, LENA  
APPLICANT: CHENG, HWEE-MING  
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN  
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8,  
TITLE OF INVENTION: 11, 16, 31, 33 AND 56,  
TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR  
TITLE OF INVENTION: DIAGNOSTIC PURPOSES  
NUMBER OF SEQUENCES: 193  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MASON & ASSOCIATES, P.A.  
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500  
CITY: CLEARWATER  
STATE: FLORIDA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 3.0  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/934,915  
FILING DATE: 22-SEP-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/949,836  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LOUISE A. Foutch  
REGISTRATION NUMBER: 37,133  
REFERENCE/DOCKET NUMBER: 1946.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 813-538-3800  
TELEFAX: 813-538-3820  
TELEX:  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-934-915-84

Query Match 62.3%; Score 35.5; DB 2; Length 20;  
Best Local Similarity 46.7%; Pred. No. 5.1;  
Matches 7; Conservative 3; Mismatches 0; Indels 5; Gaps 1;

Qy 1 LLMWIT-----QCFL 10  
Db 1 LLMWIT-----QCFL 10

Db 1 LLLWITAASAFRCFI 15

RESULT 4  
US-09-183-931-44  
Sequence 44, Application US/09183931C  
Patent No. 6210886  
GENERAL INFORMATION:  
APPLICANT: Van Baren, Nicolas  
APPLICANT: Brasseur, Francis  
APPLICANT: Boon-Failleur, Thierry  
TITLE OF INVENTION: METHOD FOR DIAGNOSING LEUKEMIA BY DETERMINING  
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSORS  
FILE REFERENCE: LUD 5527.1-JEL/ES  
CURRENT APPLICATION NUMBER: US/09/183,931C  
EARLIER FILING DATE: 2000-02-28  
EARLIER APPLICATION NUMBER: US 09/018,422  
NUMBER OF SEQ ID NOS: 44  
SEQ ID NO 44  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION:  
US-09-183-931-44

Query Match 57.9%; Score 33; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLMWIT 6  
Db 4 LLMWIT 9

RESULT 5  
US-08-764-640-35  
Sequence 35, Application US/08764640  
Patent No. 5869451  
Patent No. 5869451 5837683  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Gates, Christian  
APPLICANT: Schatz, Peter J.  
APPLICANT: Balasubramanian, Palaniappan  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Hendren, Richard W.  
APPLICANT: Deprince, Randolph B.  
APPLICANT: Podduturi, Surekha  
APPLICANT: Yin, Qun  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
TITLE OF INVENTION: RECEPTOR  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,640  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-764-640-35

Query Match 54.4%; Score 31; DB 2; Length 14;  
Best Local Similarity 62.5%; Pred. No. 20;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQC 8  
|||:|:  
Db 7 LLEWLTLC 14

RESULT 6  
US-08-764-640-36  
Sequence 36, Application US/08764640  
Patent No. 5869451  
Patent No. 5869451 5837683  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Gates, Christian  
APPLICANT: Schatz, Peter J.  
APPLICANT: Balasubramanian, Palaniappan  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Hendren, Richard W.  
APPLICANT: Deprience, Randolph B.  
APPLICANT: Podduturi, Surekha  
APPLICANT: Yin, Qun  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,640  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-764-640-36

Query Match 54.4%; Score 31; DB 2; Length 14;  
Best Local Similarity 62.5%; Pred. No. 20;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQC 8  
|||:|:  
Db 7 LLEWLTLC 14

RESULT 7  
US-08-973-225-35  
Sequence 35, Application US/08973225A  
Patent No. 6083913  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Duffin, David J.  
APPLICANT: Gates, Christian  
APPLICANT: Haselden, Sherril S.  
APPLICANT: Mattheakis, Larry C.  
APPLICANT: Schatz, Peter J.  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Wrighton, Nicholas C.  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,225A  
FILING DATE: 04-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3065USW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-08-973-225-35

Query Match 54.4%; Score 31; DB 3; Length 14;  
Best Local Similarity 62.5%; Pred. No. 20;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQC 8  
|||:|:  
Db 7 LLEWLTLC 14

RESULT 8

US-08-973-225-36  
; Sequence 36, Application US/08973225A  
; Patent No. 6083913  
; GENERAL INFORMATION:  
; APPLICANT: Dower, William J.  
; Barrett, Ronald W.  
; Cwirla, Steven E.  
; Duffin, David J.  
; Gates, Christian  
; Haselden, Sherril S.  
; Matheakis, Larry C.  
; Schatz, Peter J.  
; Wagstrom, Christopher R.  
; Wrighton, Nicholas C.

TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
THROMBOPOIETIN RECEPTOR

NUMBER OF SEQUENCES: 232

CORRESPONDENCE ADDRESS:

ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/973,225A

FILING DATE: 04-Dec-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hrubiec, Robert T.

REGISTRATION NUMBER: 36,392

REFERENCE/DOCKET NUMBER: PK3065USW

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-248-1000

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 36:

US-08-973-225-36

Query Match 54.4%; Score 31; DB 3; Length 14;  
Best Local Similarity 62.5%; Pred. NO. 20;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQC 8  
||| | | |  
DB 7 LLEWLTLC 14

RESULT 9  
US-09-244-298A-35  
; Sequence 35, Application US/09244298A  
; Patent No. 6121238  
; GENERAL INFORMATION:  
; APPLICANT: Dower, William J.  
; Barrett, Ronald W.  
; Cwirla, Steven E.  
; Gates, Christian  
; Schatz, Peter J.  
; Balasubramanian, Palaniappan  
; Wagstrom, Christopher R.  
; Hendren, Richard W.  
; DepPrince, Randolph B.  
; Podduturi, Surekha

; APPLICANT: Yin, Qun  
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
; RECEPTOR  
; NUMBER OF SEQUENCES: 244  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Glaxo Wellcome  
; STREET: Five Moore Drive, P.O. Box 13398  
; CITY: Research Triangle Park  
; STATE: NC  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/244,298A  
; FILING DATE: 11-DEC-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hrubiec, Robert T.  
; REGISTRATION NUMBER: 36,392  
; REFERENCE/DOCKET NUMBER: PK3281  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-248-1000  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-244-298A-35

Query Match 54.4%; Score 31; DB 4; Length 14;  
Best Local Similarity 62.5%; Pred. NO. 20;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQC 8  
||| | | |  
DB 7 LLEWLTLC 14

RESULT 10  
US-09-244-298A-36  
; Sequence 36, Application US/09244298A  
; Patent No. 6121238  
; GENERAL INFORMATION:  
; APPLICANT: Dower, William J.  
; Barrett, Ronald W.  
; Cwirla, Steven E.  
; Gates, Christian  
; Schatz, Peter J.  
; Balasubramanian, Palaniappan  
; Wagstrom, Christopher R.  
; Hendren, Richard W.  
; DepPrince, Randolph B.  
; Podduturi, Surekha  
; Yin, Qun  
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
; RECEPTOR  
; NUMBER OF SEQUENCES: 244  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Glaxo Wellcome  
; STREET: Five Moore Drive, P.O. Box 13398  
; CITY: Research Triangle Park  
; STATE: NC  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/244,298A
; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-244-298A-36

Query Match 54.4%; Score 31; DB 4; Length 14;
Best Local Similarity 62.5%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQC 8
Db 7 LLENLTLC 14

RESULT 11
PCT-US91-02942-14
; Sequence 14, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADAIR, JOHN R
; APPLICANT: ATHWAL, DILJEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02942
; FILING DATE: 19910429
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9009549.8
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAM L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 1011.0586600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

PCT-US91-02942-14

Query Match 49.1%; Score 28; DB 5; Length 16;
Best Local Similarity 66.7%; Pred. No. 69;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMWIT 6
Db 11 LLLWLTL 16

RESULT 12
US-08-470-419-2
; Sequence 2, Application US/08470419
; Patent No. 5866320
; GENERAL INFORMATION:
; APPLICANT: ROVINSKI, Benjamin
; APPLICANT: CAO, Shi-Xian
; APPLICANT: YAO, Fei-Long
; APPLICANT: PERSSON, Roy
; APPLICANT: KLEIN, Michel H
; TITLE OF INVENTION: ANTIGENICALLY-MARKED NON-INFECTIOUS
; RETROVIRUS-LIKE PARTICLES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,419
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290,105
; FILING DATE: August 15, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-385 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-470-419-2

Query Match 48.2%; Score 27.5; DB 2; Length 24;
Best Local Similarity 41.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

QY 2 LLMWIT---QCFL 10
Db 2 ILWISFAISCFL 13

RESULT 13
US-08-761-828-2
; Sequence 2, Application US/08761828
; Patent No. 5879925

```

;; GENERAL INFORMATION:  
;; APPLICANT: ROVINSKI, Benjamin  
;; APPLICANT: CAO, Shi-Xian  
;; APPLICANT: YAO, Fei-Long  
;; APPLICANT: PERSSON, Roy  
;; APPLICANT: KLEIN, Michel H  
;; TITLE OF INVENTION: ANTIGENICALLY-MARKED NON-INFECTIONOUS RETROVIRUS-LIKE PARTICLES  
;; NUMBER OF SEQUENCES: 26  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sim & McBurney  
;; STREET: 6TH Floor, 330 University Avenue  
;; CITY: Toronto  
;; STATE: Ontario  
;; COUNTRY: Canada  
;; ZIP: M5G 1R7  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/761.828  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/290,105  
;; FILING DATE: 15-AUG-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: STEWART, Michael I  
;; REGISTRATION NUMBER: 24,973  
;; REFERENCE/DOCKET NUMBER: 1038-655 MIS:jb  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (416) 595-1155  
;; TELEFAX: (416) 595-1163  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 24 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-08-761-828-2

Query Match 48.2%; Score 27.5; DB 2; Length 24;  
Best Local Similarity 41.7%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

Qy 2 LMWIT---QCFL 10  
Db 2 ILWISFAISCFL 13

RESULT 14  
US-08-290-105-2  
;; Sequence 2, Application US/08290105  
;; Patent No. 595342  
;; GENERAL INFORMATION:  
;; APPLICANT: ROVINSKI, Benjamin  
;; APPLICANT: CAO, Shi-Xian  
;; APPLICANT: YAO, Fei-Long  
;; APPLICANT: PERSSON, Roy  
;; APPLICANT: KLEIN, Michel H  
;; TITLE OF INVENTION: ANTIGENICALLY-MARKED NON-INFECTIONOUS  
;; RETROVIRUS-LIKE PARTICLES  
;; NUMBER OF SEQUENCES: 26  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sim & McBurney  
;; STREET: Suite 701, 330 University Avenue  
;; CITY: Toronto  
;; STATE: Ontario  
;; COUNTRY: Canada  
;; ZIP: M5G 1R7  
;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/290,105  
;; FILING DATE: August 15, 1994  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: STEWART, Michael I  
;; REGISTRATION NUMBER: 24,973  
;; REFERENCE/DOCKET NUMBER: 1038-385 MIS:jb  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (416) 595-1155  
;; TELEFAX: (416) 595-1163  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 24 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-08-290-105-2

Query Match 48.2%; Score 27.5; DB 2; Length 24;  
Best Local Similarity 41.7%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

Qy 2 LMWIT---QCFL 10  
Db 2 ILWISFAISCFL 13

RESULT 15  
US-08-776-949-2  
;; Sequence 2, Application US/08776949  
;; Patent No. 6025125  
;; GENERAL INFORMATION:  
;; APPLICANT: ROVINSKI, Benjamin  
;; APPLICANT: CAO, Shi-Xian  
;; APPLICANT: YAO, Fei-Long  
;; APPLICANT: PERSSON, Roy  
;; APPLICANT: KLEIN, Michel H  
;; TITLE OF INVENTION: ANTIGENICALLY-MARKED NON-INFECTIONOUS  
;; RETROVIRUS-LIKE PARTICLES  
;; NUMBER OF SEQUENCES: 31  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sim & McBurney  
;; STREET: 6th Floor, 330 University Avenue  
;; CITY: Toronto  
;; STATE: Ontario  
;; COUNTRY: Canada  
;; ZIP: M5G 1R7  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/776,949  
;; FILING DATE: 02-JUN-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Stewart, Michael I  
;; REGISTRATION NUMBER: 24,973  
;; REFERENCE/DOCKET NUMBER: 1038-673 MIS:jb  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (416) 595-1155  
;; TELEFAX: (416) 595-1163  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 24 amino acids  
;; TYPE: amino acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-776-949-2

Query Match 48.2%; Score 27.5; DB 3; Length 24;  
Best Local Similarity 41.7%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

QY 2 LMWIT---QCFL 10  
::||:  
Db 2 ILWISPAISCFL 13

Search completed: May 8, 2001, 15:05:10  
Job time: 106 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:08:18 ; Search time 68.24 Seconds  
(Without alignments)  
10.071 Million cell updates/sec

Title: US-09-165-546A-7  
Perfect score: 57  
Sequence: 1 LLMWITQCF 10

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 4929

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.67.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	25	43.9	24	2 F24406	unspecific monooxy
2	24	42.1	23	2 S31210	collagen alpha 2(I)
3	23	40.4	19	2 S60633	H+-transporting AT
4	23	40.4	23	2 PH1361	Ig heavy chain DJ
5	23	40.4	25	2 B36934	orf3 3' of mada -
6	22	38.6	10	2 P70289	Ig heavy chain CRD
7	22	38.6	12	2 J50424	urotensin II-B pep
8	22	38.6	12	2 I58273	thyroglobulin - ra
9	22	38.6	18	2 S19914	choline O-acetyltr
10	22	38.6	22	2 S07817	spike glycoprotein
11	21	36.8	20	2 B33761	actin - Acanthamo
12	21	36.8	23	2 S65141	nucleotide triphos
13	21	36.8	24	2 E53613	plectoxin XIII - s
14	21	36.8	24	2 A36912	hypothetical prote
15	20	35.1	7	2 S09652	hypothetical prote
16	20	35.1	7	2 PX0008	glucuronosyltransf
17	20	35.1	9	2 S07241	litorin - Rohde's
18	20	35.1	10	2 T17075	cytochrome-c oxida
19	20	35.1	13	2 P00445	urotensin II - lau
20	20	35.1	17	2 S05917	chorion class A pr
21	20	35.1	20	2 S38763	S-adenosyl-L-methi
22	20	35.1	21	2 PQ0119	dystrophin - pig (
23	20	35.1	22	2 A35418	brain natriuretic
24	20	35.1	22	2 S06487	glycine cleavage s
25	20	35.1	22	2 S42567	cytochrome-b5 redu
26	20	35.1	22	2 S01808	hemoglobin AII - t
27	19	33.3	8	2 S19288	acylase - Kluyvera
28	19	33.3	13	2 B19434	probable sex-speci
29	19	33.3	17	2 B23219	high-cysteine chor

30	19	33.3	17	2 A59069	excretory gland al
31	19	33.3	19	2 S71871	glutathione transf
32	19	33.3	20	2 S77981	cytochrome-c oxida
33	19	33.3	20	2 I49423	cytotoxic T-lympho
34	19	33.3	20	2 A37111	ribulose-bisphosph
35	19	33.3	20	2 C44920	2-halobenzoate 1,2
36	19	33.3	21	2 C38837	T-cell receptor be
37	19	33.3	21	2 S09517	prolamin - sorghum
38	19	33.3	21	2 S09516	prolamin - gamma gr
39	19	33.3	21	2 S39455	34k ribonucleoprot
40	19	33.3	23	1 S20453	pyroloquinoline q
41	19	33.3	24	2 S77982	cytochrome-c oxida
42	19	33.3	24	2 S22176	pop protein - Bac
43	19	33.3	25	2 A33058	B52b delta-like to
44	19	33.3	25	2 S10797	hypothetical prote
45	18.5	32.5	19	2 PH1353	Ig heavy chain DJ

ALIGNMENTS

RESULT 1  
F24406  
unspecific monooxygenase (EC 1.14.14.1), 53k, 3-methylcholanthrene-inducible, hepatic.  
N:Alternate names: cytochrome P450  
C:Species: *Canis porcellus* (guinea pig)  
C:Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 05-Mar-1999  
C:Accession: F24406  
R:Cheng, K.C.; Park, S.S.; Krutzsch, H.C.; Grantham, P.H.; Gelboin, H.V.; Friedman, F  
Biochemistry 25, 2397-2402, 1986  
A:Title: Amino-terminal sequence and structure of monoclonal antibody immunopurified  
A:Reference number: A24406; MUID:86243357  
A:Accession: F24406  
A:Molecule type: protein  
A:Residues: 1-24 <CHE>  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: electron transfer; heme; liver; monooxygenase; oxidoreductase; transmembr

Query Match 43.9% Score 25; DB 2; Length 24;  
Best Local Similarity 33.3% Pred. No. 3.5e+02;  
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 LLMWITQCF 9  
Db 16 LVLWVXIF 24  
RESULT 2  
S31210  
collagen alpha 2(I) chain precursor - rat (fragment)  
C:Species: *Rattus norvegicus* (Norway rat)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 13-Aug-1999  
R:Guenette, D.K.; Ritzenthaler, J.D.; Foley, J.; Jackson, J.D.; Smith, B.D.  
Biochem. J. 283, 699-703, 1992  
A:Title: DNA methylation inhibits transcription of procollagen alpha-2(I) promoters.  
A:Reference number: S31210; MUID:92272666  
A:Accession: S31210  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-23 <GUE>  
A:Cross-references: EMBL:X66209; NID:g55974; PTDN:CAA46960.1; PID:g55975  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxy-terminal homolo  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix

Query Match 42.1% Score 24; DB 2; Length 23;  
Best Local Similarity 50.0% Pred. No. 5.1e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 LLMWITQCF 8  
Db 11 LVLWVXIF 24

```

Db      11 LLLAVTSC 18

RESULT  3
S60633
H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - brine shrimp mitochondrion (strain
C:Species: mitochondrion Artemia sp. (brine shrimp)
A:Variety: strain La Mata
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-Dec-1999
C:Accession: S60633
R:Perez, M.L.; Valverde, J.R.; Batuecas, B.; Amat, F.; Marco, R.; Garesse, R.
J. Mol. Evol. 38, 156-168, 1994
A:Title: Speciation in the Artemia genus: mitochondrial DNA analysis of bisexual and par
A:Reference number: S60624; MUID:94223692
A:Accession: S60633
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-19 <PER>
A:Cross-references: EMBL:X67263; NID:gl1210; PIDN:CAA47685.1; PID:gl1211
A:Experimental source: strain La Mata
A:Note: the source is designated as Artemia parthenogenetica
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
C:Genetics:
A:Gene: ATP8
A:Genome: mitochondrion
A:Genetic code: GCC4
C:Superfamily: H+-transporting ATP synthase protein 8
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion; ox
Query Match      40.4%; Score 23; DB 2; Length 19;
Best Local Similarity 62.5%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      2 LLMWITQCF 9
      1 11111
Db      7 LPWITIFF 14

RESULT  4
PH1361
Ig heavy chain DJ region (clone CI78-122B) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1361
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A:Reference number: PH1302; MUID:93094761
A:Accession: PH1361
A:Molecule type: DNA
A:Residues: 1-23 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match      40.4%; Score 23; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 7.6e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy      1 LLMWITQCF 9
      1 11111
Db      1 VLLWFGELF 9

RESULT  5
B36934
orf3 3' of madA - Thiobacillus versutus (fragment)
C:Species: Thiobacillus versutus
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: B36934
R:Huitema, F.; van Beunemen, J.; van Driessche, G.; Duine, J.A.; Canters, G.W.
J. Bacteriol. 175, 6254-6259, 1993
A:Title: Cloning and sequencing of the gene coding for the large subunit of methylami
A:Reference number: A36934; MUID:94012487
A:Accession: B36934
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-25 <HUI>
A:Experimental source: ATCC 25364T
A:Note: sequence extracted from NCBI backbone (NCBIN:138060, NCBIP:138062)

Query Match      40.4%; Score 23; DB 2; Length 25;
Best Local Similarity 40.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy      1 LLMWITQCF 10
      1 11111
Db      9 LIHWALRSEL 18

RESULT  6
PT0289
Ig heavy chain CRD3 region (clone 4-109) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0289
R:Iyamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A:Reference number: PT0222; MUID:91108337
A:Accession: PT0289
A:Molecule type: DNA
A:Residues: 1-10 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match      38.6%; Score 22; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy      1 LLMWIT 6
      1 11111
Db      1 LLNWIS 6

RESULT  7
JS0424
urotensin II-B peptide - white sucker
C:Species: Catostomus commersoni (white sucker)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C:Accession: JS0424
R:McMaster, D.; Lederis, K.
Peptides 4, 367-373, 1983
A:Title: Isolation and amino acid sequence of two urotensin II peptides from Catostom
A:Reference number: JS0423; MUID:84041959
A:Accession: JS0424
A:Molecule type: protein
A:Residues: 1-12 <MCM>
C:Comment: This peptide has smooth muscle-stimulating activity.
F:6-11/Disulfide bonds: #status experimental

Query Match      38.6%; Score 22; DB 2; Length 12;
Best Local Similarity 75.0%; Pred. No. 6.4e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy      6 TQCF 9
      1 11111
Db      4 TECF 7

RESULT  8
158273

```

thyroglobulin - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-Nov-1999  
 C:Accession: I58273  
 R:Musti, A.M.; Ursini, V.M.; Avvedimento, E.V.; Zimarino, V.; Di Lauro, R.  
 Nucleic Acids Res. 15, 8149-8166, 1987  
 A:Title: A cell type specific factor recognizes the rat thyroglobulin promoter.  
 A:Reference number: I58273; MUID:88040446  
 A:Accession: I58273  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-12 <RES>  
 A:Cross-references: EMBL:X06162; NID:g57368; PIDN:CAA29519.1; PID:g57369

Query Match 38.6%; Score 22; DB 2; Length 12;  
 Best Local Similarity 33.3%; Pred. No. 6.4e+02;  
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMWIT 6  
 : : : :  
 Db 4 LVLWVS 9

RESULT 9  
 S19914  
 choline O-acetyltransferase (EC 2.3.1.6) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 05-May-2000  
 C:Accession: S19914  
 R:Toussaint, J.L.; Bausero, P.; Stricker, C.; Geoffroy, V.; Simoni, P.; Kempf, J.; Schmid  
 submitted to the EMBL Data Library, March 1992  
 A:Description: Human choline acetyltransferase gene: analysis of transcription control  
 A:Reference number: S19914  
 A:Accession: S19914  
 A:Molecule type: DNA  
 A:Residues: 1-18 <FOU>  
 A:Cross-references: EMBL:X65023  
 C:Keywords: acyltransferase; coenzyme A

Query Match 38.6%; Score 22; DB 2; Length 18;  
 Best Local Similarity 60.0%; Pred. No. 9.2e+02;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 MWITQ 7  
 : : : :  
 Db 7 LWCTQ 11

RESULT 10  
 S07817  
 spike glycoprotein precursor - rabies virus (fragment)  
 N:Alternate names: transmembrane glycoprotein G  
 C:Species: rabies virus  
 C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 20-Sep-1999  
 C:Accession: S07817  
 R:Poch, O.; Tordo, N.; Keith, G.  
 Biochimie 70, 1019-1029, 1988  
 A:Title: Sequence of the 386 3' nucleotides of the genome of the AV01 strain rabies virus  
 A:Reference number: S07813; MUID:89150295  
 A:Accession: S07817  
 A:Molecule type: genomic RNA  
 A:Residues: 1-22 <POC>  
 A:Cross-references: EMBL:X13357; NID:g61809; PIDN:CAA31737.1; PID:g61814  
 C:Genetics:  
 A:Gene: G  
 C:Superfamily: rabies virus spike glycoprotein  
 C:Keywords: glycoprotein; transmembrane protein

Query Match 38.6%; Score 22; DB 2; Length 22;  
 Best Local Similarity 44.4%; Pred. No. 1.1e+03;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 LLMWITQCF 9  
 : : : :  
 Db 10 LLLGFSLCF 18  
 RESULT 11  
 B33761  
 actin - Acanthamoeba castellanii (fragment)  
 C:Species: Acanthamoeba castellanii  
 C>Date: 30-Mar-1990 #sequence\_revision 30-Mar-1990 #text\_change 12-Apr-1995  
 C:Accession: B33761  
 R:Vandekerckhove, J.S.; Kaiser, D.A.; Pollard, T.D.  
 J. Cell Biol. 109, 619-626, 1989  
 A:Title: Acanthamoeba actin and profilin can be cross-linked between glutamic acid 36  
 A:Reference number: A33761; MUID:89340644  
 A:Accession: B33761  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <VAN>  
 C:Superfamily: actin

Query Match 36.8%; Score 21; DB 2; Length 20;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WITQ 7  
 : : : :  
 Db 1 WITK 4

RESULT 12  
 S65141  
 nucleotide triphosphatase - garden pea (fragments)  
 C:Species: Pisum sativum (garden pea)  
 C>Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
 C:Accession: S65141  
 R:Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.  
 Plant Mol. Biol. 30, 135-147, 1996  
 A:Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chroma  
 A:Reference number: S65141; MUID:96197404  
 A:Accession: S65141  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-15;16-23 <HSI>

Query Match 36.8%; Score 21; DB 2; Length 23;  
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 MWIT 6  
 : : : :  
 Db 17 LWVT 20

RESULT 13  
 E53613  
 plectoxin XIII - spider (Plectreurys tristis)  
 C:Species: Plectreurys tristis  
 C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 21-Aug-1998  
 C:Accession: E53613  
 R:Quistad, G.B.; Skinner, W.S.  
 J. Biol. Chem. 269, 11098-11101, 1994  
 A:Title: Isolation and sequencing of insecticidal peptides from the primitive hunting  
 A:Reference number: A53613; MUID:94209277  
 A:Accession: E53613  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-24 <QUI>  
 C:Superfamily: curatatoxin

Query Match 36.8%; Score 21; DB 2; Length 24;  
Best Local Similarity 40.0%; Pred. No. 1.8e+03;  
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WITOC 8  
I: |  
Db 7 WDYC 11

RESULT 14  
A36912  
hypothetical protein 1 baig-region [imported] - Eubacterium sp. (fragment)  
C:Species: Eubacterium Sp.  
C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000  
C:Accession: A36912  
R:Franklund, C.V.; Baron, S.F.; Hylemon, P.B.  
J. Bacteriol. 175, 3002-3012, 1993  
A:Title: Characterization of the baiH gene encoding a bile acid-inducible NADH:flavin ox  
A:Reference number: A36912; MUID:93259945  
A:Accession: A36912  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-24 <FRA>  
A:Experimental source: VPI 12708  
A:Note: sequence extracted from NCBI backbone (NCBIN:131883, NCBIPI:131906)

Query Match 36.8%; Score 21; DB 2; Length 24;  
Best Local Similarity 42.9%; Pred. No. 1.8e+03;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMWITQ 7  
I: | | | | |  
Db 7 LILFLTQ 13

RESULT 15  
S09652  
hypothetical protein (aacC2 3' region) - Enterobacter cloacae (fragment)  
C:Species: Enterobacter cloacae  
C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 08-Oct-1999  
C:Accession: S09652  
R:Vliegenthart, J.S.; Ketelaar-van Gaalen, P.A.G.; van de Klundert, J.A.M.  
Antimicrob. Agents Chemother. 33, 1153-1159, 1989  
A:Title: Nucleotide sequence of the aacC2 gene, a gentamicin resistance determinant inv  
A:Reference number: S09651; MUID:90024972  
A:Accession: S09652  
A:Molecule type: DNA  
A:Residues: 1-7 <VLI>  
A:Cross-references: EMBL:X51534; NID:g40878; PIDN:CAA35914.1; PID:g581034

Query Match 35.1%; Score 20; DB 2; Length 7;  
Best Local Similarity 40.0%; Pred. No. 2e+05;  
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMWI 5  
I: | | |  
Db 1 MIIWI 5

Search completed: May 8, 2001, 15:08:18  
Job time: 289 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:05 ; Search time 40.07 Seconds  
(without alignments)

8.549 Million cell updates/sec

Title: US-09-165-546A-7

Perfect score: 57

Sequence: 1 LLMWITQCFL 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	38.6	12	1 UR2B_CATCO	P04559 catostomus
2	22	38.6	12	1 UR2B_CYPCA	P04561 cyprinus ca
3	22	38.6	12	1 VGLG_RABVA	P15199 rabies viru
4	21	36.8	12	1 N040_SOYBN	P55960 glycine max
5	21	36.8	24	1 PL13_PLETR	P36987 plectreurus
6	20	35.1	9	1 LITR_PHYRO	P08946 phyllomedus
7	20	35.1	19	1 DCAM_ACACA	P34039 acanthamoeb
8	20	35.1	22	1 MLP_RANTE	P56924 rana tempor
9	19	33.3	17	1 CHH3_BOMMO	P20729 bombyx mori
10	19	33.3	20	1 COXA_THUOB	P80972 thunnus obe
11	19	33.3	23	1 PQQA_KLEPN	P27503 klebsiella
12	19	33.3	24	1 COXC_THUOB	P80973 thunnus obe
13	19	33.3	24	1 PSAC_CUCSA	P42046 cucumis sat
14	18	31.6	9	1 DNFL_LOCMI	P16339 locusta mig
15	18	31.6	11	1 RANC_RANPI	P08951 rana pipien
16	18	31.6	12	1 UR2_POLSP	P81022 polyodon.sp
17	18	31.6	13	1 N040_VICSA	P55961 vicia sativ
18	18	31.6	17	1 RANR_RANRU	P08952 rana rugosa
19	18	31.6	20	1 COGL_CHIOP	P34153 chionoecete
20	18	31.6	24	1 CLPP_HORVU	P48883 hordeum vul
21	18	31.6	25	1 Y15_BPT3	P20835 bacterioph
22	18	31.6	25	1 YCX8_ODOSI	P49834 odontella s
23	17	29.8	9	1 CONO_CONGE	P05486 conus geogr
24	17	29.8	15	1 ALLS_MANSE	P42559 manduca sex
25	17	29.8	20	1 CPXK_RHORH	P31718 rhodococcus
26	17	29.8	23	1 TRYL_RAF	P27436 rattus norv
27	16	28.1	9	1 ISOT_CYPCA	P42993 cyprinus ca
28	16	28.1	9	1 OXYT_EISFO	P42998 eisenia foe
29	16	28.1	9	1 OXYT_RAJCL	P42994 raja clav
30	16	28.1	12	1 UR2_SCYCA	P35490 scyllorhinu
31	16	28.1	16	1 LFKL_LOCMI	P20404 locusta mig
32	16	28.1	18	1 GOME_ACAGO	P82358 acanthoscur
33	16	28.1	18	1 TRN2_SCYCA	P08609 scyllorhinu

34 16 28.1 20 1 LYC\_FELCA P37155 felis silve  
35 16 28.1 24 1 LCA\_FELCA P37154 felis silve  
36 16 28.1 24 1 PQQA\_ACICA P27532 acinetobact  
37 15 26.3 5 1 UF01\_MOUSE P38639 mus musculu  
38 15 26.3 9 1 MGMT\_BOVIN P29177 bos taurus  
39 15 26.3 9 1 OXYA\_SCYCA P42996 scyllorhinu  
40 15 26.3 9 1 OXYA\_SQUAC P42999 squalus aca  
41 15 26.3 10 1 APE\_CAPGI P80474 capnocytoph  
42 15 26.3 10 1 RCA\_PINPS P81084 pinus pinas  
43 15 26.3 12 1 UR2A\_CATCO P04558 catostomus  
44 15 26.3 12 1 UR2\_GILMI P01147 gillichthys  
45 15 26.3 13 1 UVRD\_SALTY Q05311 salmonella

## ALIGNMENTS

RESULT 1

UR2B\_CATCO STANDARD; PRT; 12 AA.

AC P04559;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE UROTENSIN IIB (U-IIB) (UIIB).

OS Catostomus commersoni (White sucker).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Catostomidae; Catostomus.

OX NCBI\_TaxID=7971;

RN [1]

RP SEQUENCE.

RX MEDLINE=84041959; PubMed=6138758;

RA McMaster D., Lederis K.;

RT "Isolation and amino acid sequence of two urotensin II peptides from

RT Catostomus commersoni urophyses.";

RL Peptides 4:367-373(1983).

CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY

CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A

CC CORTICOTROPIN-RELEASING FACTOR.

CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.

DR PIR; JS0424; JS0424.

DR InterPro; IPR001483; -.

DR Pfam; PF02083; Urotensin\_II; 1.

DR PROSITE; PS00984; UROTENSIN\_II; 1.

KW Hormone.

FT DISULFID 6 11

SQ SEQUENCE 12 AA; 1437 MW; 73961BDBB879CEBB CRC64;

Query Match 38.6%; Score 22; DB 1; Length 12;

Best Local Similarity 75.0%; Pred. No. 2.7e+02;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TQCF 9

I::I

Db 4 TECF 7

RESULT 2

UR2B\_CYPCA

ID UR2B\_CYPCA STANDARD; PRT; 12 AA.

AC P04561;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE UROTENSIN II-BETA.

OS Cyprinus carpio (Common carp).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Cyprinus.

OX NCBI\_TaxID=7962;

RN [1]

```

RP SEQUENCE.
RA Muekata E., Ohtaki T., Ichikawa T., McMaster D., Lederis K.;
RL (In) Rich D.H., Gross E. (eds.);
RL Proceedings of the 7th American peptide symposium, pp.69-72,
RL Pierce Chemical Co., Rockford IL. (1981).
CC -!- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY
CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A
CC CORTICOTROPIN-RELEASING FACTOR.
CC -!- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.
DR InterPro; IPR001483;
DR Pfam; PF02083; Urotensin_II; 1.
DR PROSITE; PS00984; UROTENSIN_II; 1.
KW Hormone.
FT DISULFID 6 11
FT VARIANT 2 2 G->S.
SQ SEQUENCE 12 AA; 1407 MW; 73960A9FB879CEBB CRC64;

Query Match 38.6%; Score 22; DB 1; Length 12;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 TCQF 9
DB 4 TECF 7

RESULT 3
VGLG_RABVA
ID VGLG_RABVA STANDARD; PRT; 22 AA.
AC P15199;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE SPIKE GLYCOPROTEIN PRECURSOR (FRAGMENT).
GN G.
OS Rabies virus (strain AVO1).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=11293;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89150295; PubMed=3147698;
RA Poch O., Tordo N., Keith G.;
RT "Sequence of the 3386 3' nucleotides of the genome of the AVO1 strain
RT rabies virus: structural similarities in the protein regions involved
RT in transcription."
RL Biochimie 70:1019-1029(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13357; CAA31737.1;
DR PIR; S07817; S07817.
KW Transmembrane; Envelope protein; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 >22 SPIKE GLYCOPROTEIN.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2469 MW; BBA59CE12F3EFE85 CRC64;

Query Match 38.6%; Score 22; DB 1; Length 22;
Best Local Similarity 44.4%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLMWITQCF 9
DB 10 LLLGFSLCF 18

```

```

RESULT 4
NO40_SOYBN STANDARD; PRT; 12 AA.
AC P55960;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE EARLY NODULIN 40.
GN ENOD40.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. WILLIAMS;
RX MEDLINE=94035161; PubMed=8220464;
RA Yang W.C., Katinakis P., Hendriks P., Smolders A., de Vries F.,
RA Spee J., van Kammen A., Bisseling T., Franssen H.;
RT "Characterization of GmENOD40, a gene showing novel patterns of cell-
RT specific expression during soybean nodule development."
RL Plant J. 3:573-585(1993).
CC -!- FUNCTION: MODULATES THE ACTION OF AUXIN, AND MAY FUNCTION AS PLANT
CC GROWTH REGULATOR THAT ALTERS PHYTHORMONE RESPONSES (BY
CC SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE EARLY STAGES OF THE NODULE
CC DEVELOPMENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69154; -; NOT_ANNOTATED_CDS.
KW Nodulation.
SQ SEQUENCE 12 AA; 1391 MW; 3C6958AE78B1A733 CRC64;

Query Match 36.8%; Score 21; DB 1; Length 12;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMWIT 6
DB 3 LCWLT 7

RESULT 5
PLI3_PLETR
ID PLI3_PLETR STANDARD; PRT; 24 AA.
AC P36987;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PLECTOXIN XIII (PLT-XIII) (PLTXIII) (FRAGMENT).
OS Plectreureys tristis (Spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Haplogynae; Plectreureidae; Plectreureys.
OX NCBI_TaxID=33319;
RN [1]
RP SEQUENCE.
RX TISSUE=Venom;
RX MEDLINE=94209277; PubMed=8157635;
RA Quistad G.B., Skinner W.S.;
RT "Isolation and sequencing of insecticidal peptides from the primitive
RT hunting spider, Plectreureys tristis (Simon)."
RL J. Biol. Chem. 269:11098-11101(1994).

```

CC -!- FUNCTION: POTENT TOXIN THAT MAY PARALYZE AND/OR KILL INSECT  
 CC PESTS SUCH AS H. VIRESCENS (LEPIDOPTERA), S. EXIGUA (BEET ARMYWORM)  
 CC AND M. SEXTA (TOBACCO HORNWORM).  
 CC -!- PTM: MAY POSSESS FIVE DISULFIDE BONDS.  
 CC Toxin; Venom.  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2745 MW; F74A6889E80B0A9A CRC64;

Query Match 36.8%; Score 21; DB 1; Length 24;  
 Best Local Similarity 40.0%; Pred. No. 7.8e+02;  
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WITQC 8  
 Db 7 WWDYC 11

RESULT 6  
 LITR\_PHYRO STANDARD; PRT; 9 AA.  
 ID AC P08946;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RHODEI-LITORIN.  
 OS Phyllomedusa rohdei (Rohde's leaf frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
 OC Phyllomedusa.  
 OX NCBI\_TaxID=8394;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=85127560; PubMed=3838283;  
 RA Bara D., Erspamer G.F., Simmaco M., Bossa F., Melchiorri P.,  
 RA Erspamer V.;  
 RT "Rhodei-litorin: a new peptide from the skin of Phyllomedusa rohdei".  
 RL FEBS Lett. 182:53-56(1985).  
 CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN  
 CC FAMILY.  
 DR PIR: S07241; S07241.  
 DR InterPro: IPR000874; .  
 DR Pfam: PF02044; Bombesin; 1.  
 DR PROSITE: PS00257; BOMBESIN; 1.  
 KW Bombesin family; Amidation.  
 FT MOD\_RES 1 1  
 FT MOD\_RES 9 9  
 FT MOD\_RES 9 9  
 SQ SEQUENCE 9 AA; 1090 MW; 4ECCC1E861ADC377 CRC64;

Query Match 35.1%; Score 20; DB 1; Length 9;  
 Best Local Similarity 37.5%; Pred. No. 9.3e+04;  
 Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 MWITQCFL 10  
 Db 2 LWATGHFM 9

RESULT 7  
 DCAM\_ACACA STANDARD; PRT; 19 AA.  
 ID AC P34039;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE S-ADENOSYLMETHIONINE DECARBOXYLASE ALPHA SUBUNIT (EC 4.1.1.50)  
 DE (ADOMETDC) (FRAGMENT).  
 OS Acanthamoeba castellanii (Amoeba).  
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.  
 OX NCBI\_TaxID=5755;  
 RN [1]  
 RP SEQUENCE, AND CHARACTERIZATION.

RX MEDLINE=94029912; PubMed=8216217;  
 RA Hugo E.R., Byers T.J.;  
 RT "S-adenosyl-L-methionine decarboxylase of Acanthamoeba castellanii  
 (Neff): purification and properties.";  
 RL Biochem. J. 295:203-209(1993).  
 CC -!- FUNCTION: S-ADENOSYLMETHIONINE DECARBOXYLASE IS ESSENTIAL FOR THE  
 CC BIOSYNTHESIS OF SPERMINE AND SPERMIDINE. THE ALPHA SUBUNIT  
 CC CONTAINS THE ACTIVE SITE.  
 CC -!- CATALYTIC ACTIVITY: S-ADENOSYLMETHIONINE = (5-DEOXY-5-ADENOSYL)  
 CC (3-AMINOPROPYL)METHYLSULFONIUM SALT + CO(2).  
 CC -!- COFACTOR: REQUIRES A PYRUVYL GROUP FOR ITS ACTIVITY.  
 CC -!- PATHWAY: DECARBOXYLATION OF S-ADENOSYLMETHIONINE PROVIDES THE  
 CC AMINOPROPYL MOIETY REQUIRED FOR SPERMIDINE AND SPERMINE  
 CC BIOSYNTHESIS FROM PUTRESCINE.  
 CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED ONLY DURING EXPONENTIAL GROWTH.  
 CC -!- INDUCTION: STIMULATED BY PUTRESCINE. INHIBITED BY AROMATIC  
 CC DIAMIDINES BERENIL, PENTAMIDINE, PROPAMIDINE, HYDROXYSTILBAMIDINE,  
 CC BY ETHIDIUM BROMIDE AND METHYLGLYOXAL.  
 CC -!- SIMILARITY: BELONGS TO THE ADOMETC FAMILY.  
 DR InterPro: IPR001985;  
 DR PROSITE: PS01336; ADOMETDC; PARTIAL.  
 KW Spermidine biosynthesis; Lyase; Decarboxylase; Pyruvate.  
 FT MOD\_RES 1 1  
 FT MOD\_RES 19 19  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 2153 MW; 88B18AD9B6142AEF CRC64;

Query Match 35.1%; Score 20; DB 1; Length 19;  
 Best Local Similarity 30.0%; Pred. No. 9.3e+02;  
 Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLMWITQCFL 10  
 Db 3 MFVNTKLIL 12

RESULT 8  
 MLP\_RANTE STANDARD; PRT; 22 AA.  
 ID AC P56924;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE MELITTIN-LIKE PEPTIDE (MLP).  
 OS Rana temporaria (European common frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.  
 OX NCBI\_TaxID=8407;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin;  
 RX MEDLINE=97175050; PubMed=9022710;  
 RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,  
 RA Bara D.;  
 RT "Temporins, antimicrobial peptides from the European red frog Rana  
 temporaria".  
 RL Eur. J. Biochem. 242:788-792(1996).  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- SIMILARITY: HIGH, TO A. MILLIFERA MELITTIN.  
 KW Amphibian skin; Amidation.  
 FT MOD\_RES 22 22  
 FT MOD\_RES 22 22  
 SQ SEQUENCE 22 AA; 2313 MW; 0990320F9B0709B3 CRC64;

Query Match 35.1%; Score 20; DB 1; Length 22;  
 Best Local Similarity 28.6%; Pred. No. 1.1e+03;  
 Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQ 7  
 Db 16 IVSWVKQ 22

```

RESULT 9
CHH3_BOMMO STANDARD; PRT; 17 AA.
ID CHH3_BOMMO
AC P20729;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE CHORION CLASS HIGH-CYSTEINE HCA PROTEIN 13 PRECURSOR (HC-A.13)
DE (FRAGMENT).
OS Bombyx mori (Silk moth).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85083111; PubMed=6439880;
RA Rodakis G.C., Lecanidou R., Eickbush T.H.;
RT "Diversity in a chorion multigene family created by tandem
  duplications and a putative gene-conversion event.";
RL J. Mol. Evol. 20:265-273(1984).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
  SILK MOTH.
CC -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
  BELONG CLASSES A, CA AND HCA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01068; CAB57790.1;
DR PIR; B23219; B23219
KW Eggshell; Chorion; Repeat; Multigene family; Signal.
FT SIGNAL 1 >17
FT NON_TER 17
SQ SEQUENCE 17 AA; 1913 MW; 5E634508C5355C9C CRC64;

Query Match 33.3%; Score 19; DB 1; Length 17;
Best Local Similarity 30.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLMWITQCF 10
DB 7 LLLCVQGLI 16

RESULT 10
COXA_THUOB STANDARD; PRT; 20 AA.
ID COXA_THUOB
AC P80972;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-1 (EC 1.9.3.1) (FRAGMENT).
OS Thunus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RX TISSUE=Heart, and Liver;
RA MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
  Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and

```

```

RT liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
  OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
  4 FERRICYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
  Oxidoreductase; Heme; Mitochondrion; Inner membrane.
KW NON_TER 20
SQ SEQUENCE 20 AA; 2404 MW; 7E82E43B7157355E CRC64;

Query Match 33.3%; Score 19; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WIT 6
DB 15 WWT 17

RESULT 11
PQQA_KLEPN STANDARD; PRT; 23 AA.
ID PQQA_KLEPN
AC P27503;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE COENZYME PQQ SYNTHESIS PROTEIN A.
GN PQQA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 418;
RX MEDLINE=92212293; PubMed=1313537;
RA Meulenbergh J.J.M., Sellink E., Riegman N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqq
  operon.";
RT operson.
RL Mol. Gen. Genet. 232:284-294(1992).
CC -!- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
  BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
  SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
  ARE NECESSARY FOR THE SYNTHESIS.
CC -!- SIMILARITY: TO OTHER BACTERIAL PQQA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58778; CAA41579.1;
DR PIR; S20453; S20453.
KW PQQ.
SQ SEQUENCE 23 AA; 2764 MW; ACCB321460871C5D CRC64;

Query Match 33.3%; Score 19; DB 1; Length 23;
Best Local Similarity 37.5%; Pred. No. 1.6e+03;
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 MWITQCF 10
DB 1 MWKKPAFI 8

RESULT 12

```

COXC\_THUOB  
ID COXC\_THUOB STANDARD; PRT; 24 AA.  
AC P80373;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE VA-2 (EC 1.9.3.1) (FRAGMENT).  
OS Thunus obesus (Bigeye tuna).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;  
OC Scombridae; Thunnus.  
OX NCBI\_TaxID=8241;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Heart, and Liver;  
RX MEDLINE=97454291; PubMed=9310366;  
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,  
RA Kadenbach B.;  
RT "The subunit structure of cytochrome-c oxidase from tuna heart and liver.";  
RL Eur. J. Biochem. 248:99-103(1997).  
CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C OXIDASE. THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.  
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O + 4 FERRICYTOCHROME C.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.  
KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.  
FT NON\_TER 24  
SQ SEQUENCE 24 AA; 2903 MW; 20998FB91F22E43B CRC64;

Query Match 33.3%; Score 19; DB 1; Length 24;  
Best Local Similarity 66.7%; Pred. No. 1.7e+03;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WIT 6  
DB 15 WVT 17

RESULT 13  
PSAC\_CUCSA  
ID PSAC\_CUCSA STANDARD; PRT; 24 AA.  
AC P42046;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE PHOTOSYSTEM I IRON-SULFUR CENTER (PHOTOSYSTEM I SUBUNIT VII) (9 KDA POLYPEPTIDE) (PSI-C) (PS 1 SUBUNIT 11) (FRAGMENT).  
GS PSAC OR FRXA.  
OS Cucumis sativus (Cucumber).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Cucurbitales; Cucurbitaceae; Cucumis.  
OX NCBI\_TaxID=3659;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cotyledon;  
RX MEDLINE=91355209; PubMed=1883835;  
RA Iwasaki Y., Ishikawa H., Hibino T., Takabe T.;  
RT "Characterization of genes that encode subunits of cucumber PS I complex by N-terminal sequencing.";  
RL Biochim. Biophys. Acta 1059:141-148(1991).  
CC -!- FUNCTION: APOPROTEIN FOR THE IRON-SULFUR CENTERS FA AND FB OF THE PHOTOSYSTEM I COMPLEX. ESSENTIAL FOR PHOTOCHEMICAL ACTIVITY AND STABLE ASSEMBLY OF PSI.  
CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
CC Mendel; 7834; CUCsa:psac;1.  
DR InterPro; IPR001450; -.

Pfam: PF00037; fer4; 1.  
DR PROSITE; PS00198; 4FE4S\_FERREDOXIN; 1.  
KW Photosynthesis; Chloroplast; Photosystem I; Iron-sulfur; 4Fe-4S.  
FT INIT\_MET 0  
FT METAL 10 10  
FT METAL 13 13  
FT METAL 16 16  
FT METAL 20 20  
FT UNSURE 10 10  
FT UNSURE 13 13  
FT UNSURE 16 16  
FT UNSURE 20 20  
FT NON\_TER 24  
SQ SEQUENCE 24 AA; 2573 MW; 647776B743B5F894 CRC64;

Query Match 33.3%; Score 19; DB 1; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TQC 8  
DB 14 TQC 16

RESULT 14  
DNFL\_LOCFI  
ID DNFL\_LOCFI STANDARD; PRT; 9 AA.  
AC P16339;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE LOCUPRESSIN (DIURETIC NEUROPEPTIDE F1/F2).  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
OC Acridoidea; Acrididae; Locusta.  
OX NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Suboesophageal ganglion, and Thoracic ganglion;  
RX MEDLINE=88077077; PubMed=3689410;  
RA Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,  
RA Delaage M., Schooley D.A.;  
RT "Identification of an arginine vasopressin-like diuretic hormone from Locusta migratoria.";  
RL Biochem. Biophys. Res. Commun. 149:180-186(1987).  
CC -!- FUNCTION: DIURETIC HORMONE.  
CC -!- SUBUNIT: F2 IS AN ANTIPARALLEL DISULFIDE LINKED DIMER OF F1.  
CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
DR PIR; A29477; A29477.  
DR InterPro; IPR000981; -.  
DR Pfam; PF00220; hormone4; 1.  
DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
KW Hormone; Neuropeptide; Amidation.  
FT DISULFID 1 6  
FT DISULFID 1 1  
FT DISULFID 6 6  
FT MOD\_RES 9 9  
SQ SEQUENCE 9 AA; 976 MW; 56EB176EB451A057 CRC64;

Query Match 31.6%; Score 18; DB 1; Length 9;  
Best Local Similarity 75.0%; Pred. No. 9.3e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITQC 8  
DB 3 ITNC 6

RESULT 15  
RANC\_RANPI

ID RANC\_RANPI STANDARD; PRT; 11 AA.  
 AC P08951;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RANATENSIN-C.  
 OS Rana pipiens (Northern leopard frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.  
 OX NCBI\_TaxID=8404;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=84131098; PubMed=6141890;  
 RA Nakajima T.;  
 RL Unpublished results, cited by:  
 RL Erspaner V., Erspaner G.F., Mazzanti G., Endean R.;  
 RL Comp. Biochem. Physiol. 77C:99-108(1984).  
 CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN  
 CC FAMILY.  
 DR InterPro: IPR000874; .  
 DR Pfam: PF02044; Bombesin; 1.  
 DR PROSITE: PS00257; BOMBESIN; 1.  
 KW Bombesin family; Amidation.  
 FT MOD\_RES 11  
 FT 11  
 SQ SEQUENCE 11 AA; 1304 MW; D6C9885A61ADC366 CRC64;

Query Match 31.6%; Score 18; DB 1; Length 11;  
 Best Local Similarity 42.9%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 WITOCFL 10  
 | | |  
 Db 5 WATGHFM 11

Search completed: May 8, 2001, 15:14:05  
 Job time: 526 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:16:04 ; Search time 114.89 Seconds  
(without alignments)  
10.202 Million cell updates/sec

Title: US-09-165-546A-7  
Perfect score: 57  
Sequence: 1 LLMWITQCFL 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL\_15.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_unclassified.\*  
13: sp\_vertebrate.\*  
14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	52.6	11	Q9UE69	Q9ue69 homo sapien
2	27	47.4	20	Q36127	Q36127 thyrophigus
3	25	43.9	23	Q9NZ62	Q9nz62 homo sapien
4	24	42.1	18	Q9UJ23	Q9uj23 homo sapien
5	24	42.1	21	Q9SLC0	Q9slc0 porphyromon
6	24	42.1	11	Q83966	Q83966 influenza a
7	23	40.4	11	Q84073	Q84073 influenza a
8	23	40.4	16	Q62810	Q62810 equus caball
9	23	40.4	17	Q9QUY6	Q9quy6 rattus sp.
10	23	40.4	19	Q31687	Q31687 artemia par
11	23	40.4	23	Q9NZ61	Q9nz61 homo sapien
12	23	40.4	24	Q9UE18	Q9ue18 homo sapien
13	22	38.6	11	Q9UCR1	Q9ucr1 homo sapien
14	22	38.6	12	Q63579	Q63579 rattus norv
15	22	38.6	17	P97758	P97758 mus musculus
16	22	38.6	18	Q9ZY75	Q9zy75 cenocoelius
17	22	38.6	19	O76047	O76047 mus sapien
18	22	38.6	20	P97907	P97907 mus musculus
19	22	38.6	22	Q9W6D7	Q9w6d7 gallus gall

20	22	38.6	24	8	Q9ZY84	Q9zy84 melissodes
21	21.5	37.7	15	2	Q9R568	Q9r568 nitrosomona
22	21.5	37.7	16	9	O21922	O21922 streptococc
23	21	36.8	8	2	O85406	O85406 coxiella bu
24	21	36.8	8	4	O15890	O15890 homo sapien
25	21	36.8	13	2	O49230	O49230 mycoplasma
26	21	36.8	18	4	Q9UE78	Q9uet8 homo sapien
27	21	36.8	21	14	O84058	O84058 influenza a
28	21	36.8	22	4	Q9UN58	Q9uns8 homo sapien
29	21	36.8	22	9	Q38287	Q38287 lactococcus
30	21	36.8	23	13	Q9PS32	Q9ps32 gallus gall
31	21	36.8	24	7	Q9MY61	Q9my61 homo sapien
32	21	36.8	25	14	Q9IER9	Q9ier9 papillomavi
33	20	35.1	8	5	Q9VRD2	Q9vrd2 drosophila
34	20	35.1	10	8	O79912	O79912 chamaeleo f
35	20	35.1	13	2	Q9R3R3	Q9r3r3 borrelia bu
36	20	35.1	13	11	Q9OW45	Q9ow45 rattus sp.
37	20	35.1	17	2	Q9R9C3	Q9r9c3 borrelia bu
38	20	35.1	17	5	Q9TVX7	Q9tvx7 bombyx mori
39	20	35.1	17	14	O89807	O89807 epiphyas po
40	20	35.1	19	4	Q9UJY6	Q9ujy6 homo sapien
41	20	35.1	21	2	O9L4S9	O9l4s9 salmonella
42	20	35.1	23	4	Q9UD22	Q9udz2 homo sapien
43	20	35.1	25	2	Q9ZEW6	Q9zew6 klebsiella
44	20	35.1	25	5	Q9NSB6	Q9n5b6 caenorhabdi
45	20	35.1	25	9	Q9T139	Q9t139 bacterioph

#### ALIGNMENTS

#### RESULT 1

Q9UE69 PRELIMINARY; PRT; 11 AA.  
AC Q9UE69;  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DE 5HT3 SEROTONIN RECEPTOR (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bedford F.K., Taylor S., Julius D., Ingraham H.A.;  
RT "Expression of the 5HT3 serotonin receptor gene in neuronal cells is  
RT regulated via novel NF-1 mediated complexes."  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U73443; AAB71736.1; -.  
KW Receptor.  
FT NON\_TER.  
SQ SEQUENCE 11 AA; 1286 MW; 28F50414C6C6C2D3 CRC64;

Query Match 52.6%; Score 30; DB 4; Length 11;  
Best Local Similarity 40.0%; Pred. No. 31;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLMWITQCFL 10  
:|:|:  
Db 1 MLLWVQQALL 10

#### RESULT 2

Q36127 PRELIMINARY; PRT; 20 AA.  
ID Q36127;  
AC Q36127;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-MAY-1999 (TREMREL. 10, Last annotation update)  
DE NADH DEHYDROGENASE SUBUNIT 1 (FRAGMENT).  
GN ND1.

OS Thyrophysus sp.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Diplopoda;  
 OC Helminthomorpha; Julida; Julidae; Thyrophysus.  
 OX NCBI\_TaxID=41451;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95327187; PubMed=7603565;  
 RA Boore J.L., Collins T.M., Stanton D., Daehler L.L., Brown W.M.;  
 RT "Deducing the pattern of arthropod phylogeny from mitochondrial DNA  
 RT rearrangements";  
 RL Nature 376:163-165(1995).  
 DR EMBL: U29704; AAA82253.1; -  
 KW Mitochondrion.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2483 MW; 9E4682A1056F7774 CRC64;  
 Query Match 47.48; Score 27; DB 8; Length 20;  
 Best Local Similarity 80.08; Pred. No. 1.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLMWI 5  
 Db 5 LLMWV 9  
 RESULT 3  
 Q9NZ62  
 ID Q9NZ62 PRELIMINARY; PRT; 23 AA.  
 AC Q9NZ62;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE ATP7B (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RA Wu Z.Y., Wang N., Mu Rong S.X.;  
 RT "Another hot point mutation of Wilson disease gene in  
 RT Chinese: exon12";  
 RL Chung-Hua I Hsueh Tsa Chih 79:422-423(1999).  
 DR EMBL: AF218766; AAF34875.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 13 13 M -> T.  
 FT NON\_TER 23 23  
 SQ SEQUENCE 23 AA; 2758 MW; 10DCAF3B5A40804A CRC64;  
 Query Match 43.98; Score 25; DB 4; Length 23;  
 Best Local Similarity 40.08; Pred. No. 4.7e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 LLMWITQCF 10  
 Db 14 LVVWIVVFI 23  
 RESULT 4  
 Q9UJ23  
 ID Q9UJ23 PRELIMINARY; PRT; 18 AA.  
 AC Q9UJ23;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE OVIDUCT GLYCOPROTEIN (FRAGMENT).

GN OGP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lee K.F., Kwok K.L., Agarwal A., Lee Y.L.;  
 RT "Human oviduct glycoprotein promoter sequence";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF189710; AAF01065.1; -  
 FT NON\_TER 18 18  
 FT NON\_TER 18 18  
 SQ SEQUENCE 18 AA; 2201 MW; 5E8FD91EA210E516 CRC64;  
 Query Match 42.18; Score 24; DB 4; Length 18;  
 Best Local Similarity 60.08; Pred. No. 5.8e+02;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 LLMWI 5  
 Db 4 LLMWV 8  
 RESULT 5  
 Q9SICO  
 ID Q9SICO PRELIMINARY; PRT; 21 AA.  
 AC Q9SICO;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE TRANSPOSASE (FRAGMENT).  
 OS Porphyromonas gingivalis (Bacteroides gingivalis).  
 OC Bacteria; CFB group; Bacteroidaceae; Porphyromonas.  
 OX NCBI\_TaxID=837;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=W83;  
 RA Sawada K., Koikeguchi S., Hongyo H., Sawada S., Miyamoto M., Maeda H.,  
 RA Nishimura F., Takashiba S., Murayama Y.;  
 RT "Identification by subtractive hybridization of a novel insertion  
 RT sequence specific for virulent strains of Porphyromonas gingivalis";  
 RL Infect. Immun. 67:5621-5625(1999).  
 DR EMBL: AB011547; BAA83478.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 1 1  
 SQ SEQUENCE 21 AA; 2543 MW; 08D3F838FEB3E6AF CRC64;  
 Query Match 42.18; Score 24; DB 2; Length 21;  
 Best Local Similarity 33.38; Pred. No. 6.6e+02;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 LLMWITQCF 9  
 Db 2 LMOWLNEPF 10  
 RESULT 6  
 Q83966  
 ID Q83966 PRELIMINARY; PRT; 21 AA.  
 AC Q83966;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE INFLUENZA A/BH/35 (H1N1), NEURAMINIDASE (SEG 6) RNA (FRAGMENT).  
 OS Influenza A virus.  
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 OC Influenza virus A and B group.  
 OX NCBI\_TaxID=11320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82201010; PubMed=7080440;  
 RA Blok J., Air G.M.;

RT \*Block deletions in the neuraminidase genes from some influenza A  
RL viruses of the N1 subtype.\*;  
RL Virology 118:229-234(1982).  
DR EMBL; J02563; AAA43388.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2448 MW; 13C6CFB2007E98EF CRC64;

Query Match 42.1%; Score 24; DB 14; Length 21;  
Best Local Similarity 42.9%; Pred. No. 6.6e+02;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 WITQCF 10  
DB 10 WINQYI 16

RESULT 7  
Q84073 ID Q84073 PRELIMINARY; PRT; 11 AA.  
AC Q84073;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE INFLUENZA A/FPV/ROSTOCK/34 (H7N1), POLYMERASE 3  
DE (SEG 3), 3' END OF VRNA (INITIATOR REGION FOR PROTEIN CODING)  
DE (FRAGMENT).  
OS Influenza A virus.  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group.  
OX NCBI\_TaxID=11320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=80034428; Pubmed=493121;  
RA Robertson J.S.;  
RT "5' and 3' terminal nucleotide sequences of the RNA genome segments of  
influenza virus.\*;  
RL Nucleic Acids Res. 6:3745-3757(1979).  
DR EMBL; J02123; AAA43612.1; -.  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1400 MW; CC2007F7A6C412C9 CRC64;

Query Match 40.4%; Score 23; DB 14; Length 11;  
Best Local Similarity 50.0%; Pred. No. 5.7e+02;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WITQCF 9  
DB 4 FVRQCF 9

RESULT 8  
O62810 ID O62810 PRELIMINARY; PRT; 16 AA.  
AC O62810;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-AUG-1998 (Tremblrel. 07, Last annotation update)  
DE PREGNANCY-ASSOCIATED GLYCOPROTEIN (FRAGMENT).  
OS Equus caballus (Horse).  
GN PAG.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Green J.A., Xie S., Roberts R.M.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF061188; AAC15768.1; -.  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1830 MW; 733B85DC7AAC213F CRC64;

Query Match 40.4%; Score 23; DB 6; Length 16;  
Best Local Similarity 37.5%; Pred. No. 7.9e+02;  
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMWITQCF 9  
DB 9 LVTLSECF 16

RESULT 9  
Q9QUY6 ID Q9QUY6 PRELIMINARY; PRT; 17 AA.  
AC Q9QUY6;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
DE OLIGODENDROCYTE-SPECIFIC UDP-GALACTOSE: CERAMIDE  
DE GALACTOSYLTRANSFERASE (FRAGMENT).  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96085162; Pubmed=8521863;  
RA Schulte S., Stoffel W.;  
RT "UDP galactose:ceramide galactosyltransferase and glutamate/aspartate  
transporter. Copurification, separation and characterization of the  
two glycoproteins.\*;  
RL Eur. J. Biochem. 233:947-953(1995).  
SQ SEQUENCE 17 AA; 2017 MW; 0F92A645FFD7F828 CRC64;

Query Match 40.4%; Score 23; DB 11; Length 17;  
Best Local Similarity 42.9%; Pred. No. 8.3e+02;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LLMWITQ 7  
DB 1 LIOWLPQ 7

RESULT 10  
Q31687 ID Q31687 PRELIMINARY; PRT; 19 AA.  
AC Q31687;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE ATPASE SUBUNIT 8 (FRAGMENT).  
GN ATP8.  
OS Artemia parthenogenetica.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;  
OC Artemiidae; Artemia.  
OX NCBI\_TaxID=6663;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LA MATA;  
RX MEDLINE=94223692; Pubmed=8169960;  
RA Perez M.L., Valverde J.R., Batuecas B., Amat F., Marco R., Garesse R.;  
RT "Speciation in the Artemia genus: mitochondrial DNA analysis of  
bisexual and parthenogenetic brine shrimps.\*;  
RL J. Mol. Evol. 38:156-168(1994).  
DR EMBL; X67263; CAA47685.1; -.  
KW Mitochondrion.  
FT NON\_TER 1 1  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2246 MW; 41922AD313E087E3 CRC64;

Query Match 40.4%; Score 23; DB 8; Length 19;  
 Best Local Similarity 62.5%; Pred. No. 9.2e+02;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LWMITQCF 9  
 I:| | | |  
 Db 7 LPWITIFF 14

RESULT 11  
 Q9NZ61 PRELIMINARY; PRT; 23 AA.  
 AC Q9NZ61:  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE ATP7B (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wu Z.Y., Wang N., Mu Rong S.X.;  
 RT "Another hot point mutation of Wilson disease gene in  
 Chinese: exon12.";  
 RL Chung-Hua I Hsueh Tsa Chih 79:422-423(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Wu Z.Y., Wang N., Mu Rong S.X.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF18767; AAF34876.1; -.  
 FT NON\_TER 1 1  
 FT VARIANT 21 21 D -> G.  
 FT NON\_TER 23 23  
 FT NON\_TER 23 23  
 SQ SEQUENCE 23 AA; 2743 MW; 10DCA8665A40804A CRC64;

Query Match 40.4%; Score 23; DB 4; Length 23;  
 Best Local Similarity 40.0%; Pred. No. 1.1e+03;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LLMWITQCF 10  
 I:| | | |  
 Db 14 LVWVIVDFI 23

RESULT 12  
 Q9UE18 PRELIMINARY; PRT; 24 AA.  
 AC Q9UE18:  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE LAMININ ALPHA 4 (FRAGMENT).  
 GN LAMA4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Richards A.J., Luccarino C., Pope F.M.;  
 RT "The structural organisation of LAMA 4 the gene encoding laminin alpha  
 4.";  
 RL Eur. J. Biochem. 248:15-23(1997).  
 DR EMBL; Y14240; CAA74636.1; -.  
 FT NON\_TER 24 24  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2676 MW; 5358D6159B7A7652 CRC64;

Query Match 40.4%; Score 23; DB 4; Length 24;  
 Best Local Similarity 42.9%; Pred. No. 1.1e+03;

Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LWMITQCF 8  
 I:| | |  
 Db 15 LLWSAAC 21

RESULT 13  
 Q9UCR1 PRELIMINARY; PRT; 11 AA.  
 AC Q9UCR1:  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE AUTOTAXIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92129337; PubMed=1733949;  
 RA Stracke M.L., Krutzsch H.C., Unsworth E.J., Ciocco A., Ciocco V.,  
 RA Schiffmann E., Liotta L.A.;  
 RT "Identification, purification, and partial sequence analysis of  
 autotaxin, a novel motility-stimulating protein.";  
 RL J. Biol. Chem. 267:2524-2529(1992).  
 SQ SEQUENCE 11 AA; 1171 MW; 2723615AA0437737 CRC64;

Query Match 38.6%; Score 22; DB 4; Length 11;  
 Best Local Similarity 75.0%; Pred. No. 8.6e+02;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 MWIT 6  
 I:| | |  
 Db 5 LWIT 8

RESULT 14  
 Q63579 PRELIMINARY; PRT; 12 AA.  
 ID Q63579:  
 AC Q63579:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE THYROGLOBULIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88040446; PubMed=3671079;  
 RA Musti A.M., Ursini V.M., Avvedimento E.V., Zimarino V., Lauro R.D.;  
 RT "A cell type specific factor recognizes the rat thyroglobulin  
 promoter.";  
 RL Nucleic Acids Res. 15:8149-8166(1987).  
 DR EMBL; X06162; CAA29519.1; -.  
 FT NON\_TER 12 12  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1441 MW; D7158AD39BD37732 CRC64;

Query Match 38.6%; Score 22; DB 11; Length 12;  
 Best Local Similarity 33.3%; Pred. No. 9.2e+02;

Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLMWIT 6  
 I:| | | |  
 Db 4 LVLWVS 9

RESULT 15

P97758  
ID P97758 PRELIMINARY; PRT; 17 AA.  
AC P97758;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE DBM28 PROTEIN (FRAGMENT).  
GN H-2D.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6BY-H-2BM28;  
RA Yun T.J., Melvold R.W., Pease L.R.;  
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(0).  
DR EMBL; U83304; AAB41021.1; -  
DR INTERPRO: IPR001039; -  
DR PFAM; PF00129; MHC\_I; 1.  
FT NON\_TER 1  
FT NON\_TER 17  
SQ SEQUENCE 17 AA; 1956 MW; F3149F377C16F196 CRC64;

Query Match 38.6%; Score 22; DB 11; Length 17;  
Best Local Similarity 44.4%; Pred. No. 1.3e+03;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 LMWITOCFL 10  
| | | | |  
Db 5 LQWNYGCDL 13

Search completed: May 8, 2001, 15:16:04  
Job time: 530 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:07:06 ; Search time 113.31 Seconds  
(without alignments)  
9.081 Million cell updates/sec

Title: US-09-165-546A-8  
Perfect score: 88  
Sequence: 1 AADHRQLQLSTSSCLQQL 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 165078

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0401.\*  
1: /SID56/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID56/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID56/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID56/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID56/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID56/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID56/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID56/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID56/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID56/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID56/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID56/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID56/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID56/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID56/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID56/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID56/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID56/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID56/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID56/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID56/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID56/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	100.0	18	21 Y52435	Human tumour antig
2	48	54.5	10	20 Y06011	Human cancer antig
3	47	53.4	10	20 Y06016	Human cancer antig
4	43	48.9	9	20 Y06035	Human cancer antig
5	43	48.9	9	21 Y79748	NY-ESO-1 derived p
6	42	47.7	9	20 Y06037	Human cancer antig
7	39	44.3	9	20 Y06042	Human cancer antig
8	36	40.9	9	21 Y79754	NY-ESO-1 derived p
9	35	39.8	21	13 R29045	PLA2 inhibitory pe
10	33	37.5	23	20 Y25818	Human secreted pro
11	32	36.4	18	21 Y52440	Human tumour antig

12	29	33.0	13	20	Y45398	Immunogenic peptid
13	29	33.0	18	16	R84751	GST-phospholipase
14	29	33.0	18	20	W96889	ApoB-100 nuclear l
15	29	33.0	20	17	W11227	Modified rat obese
16	29	33.0	20	17	R92724	Immunogenic fragme
17	29	33.0	20	21	B28456	Murine OB carboxy
18	29	33.0	20	21	B28475	Murine OB carboxy
19	29	33.0	20	21	Y87734	Murine OB protein-
20	29	33.0	23	21	Y64681	Human 5' EST relat
21	28	31.8	16	21	Y59323	Tyrosine phosphata
22	28	31.8	23	21	Y51777	H. influenzae tran
23	28	31.8	23	21	Y80377	H. influenzae tran
24	28	31.8	25	21	B02059	CAP37 peptide anal
25	27	30.7	9	20	Y06034	Human cancer antig
26	27	30.7	9	20	Y06038	Human cancer antig
27	27	30.7	9	20	Y06039	Human cancer antig
28	27	30.7	10	20	Y06055	Human cancer antig
29	27	30.7	10	20	Y05999	Human cancer antig
30	27	30.7	10	20	Y05989	Human cancer antig
31	27	30.7	15	22	B65571	Immunogenic peptid
32	27	30.7	16	21	Y52670	rel-encoded oncopr
33	27	30.7	18	21	B09066	Hepatitis GB virus
34	27	30.7	19	15	R58673	Rat PACAP receptor
35	27	30.7	20	14	R36880	Guinea pig VPF C-t
36	27	30.7	20	19	W42160	T-cell epitope pep
37	27	30.7	20	21	B13623	C. trachomatis ser
38	27	30.7	23	20	Y45350	Human secreted pro
39	27	30.7	24	20	Y22024	VEGF antagonist pe
40	27	30.7	25	14	R41934	Peptide fragment o
41	27	30.7	25	18	W26803	Human Cationic Ant
42	27	30.7	25	18	W15418	CAP37 fragment, 20
43	27	30.7	25	20	W97174	Peptide derived fo
44	27	30.7	25	21	B28489	Human lipopolysacc
45	27	30.7	25	21	B02021	CAP37 peptide anal

## ALIGNMENTS

RESULT 1  
Y52435  
ID Y52435 standard; Protein; 18 AA.  
XX Y52435;  
XX AC  
XX DT  
DT 15-FEB-2000 (first entry)  
XX DE  
DE Human tumour antigen NY-ESO-1 peptide #8.

XX Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
KW T-cell; helper; stimulation; proliferation; treatment;  
KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
KW lymphoma.

XX Synthetic.  
OS Homo sapiens.  
XX PN  
PN WO9953938-A1.  
XX PD  
PD 28-OCT-1999.

XX 24-MAR-1999; 99WO-US06875.  
XX 17-APR-1998; 98US-0062422.  
PR 02-OCT-1998; 98US-0165546.  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
PI Gure A, Rittner G;  
XX WPI; 2000-038483/03.

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
 PT useful for therapeutic and diagnostic purposes -  
 XX  
 PS Claim 4; Page 22; 49pp; English.  
 XX  
 CC Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels  
 CC in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and  
 CC tumour cell lines with some degree of frequency - these included  
 CC melanoma specimens and cell lines, and breast and bladder cancer  
 CC specimens, with expression in other tumour types being sporadic.  
 CC These NY-ESO-1-derived peptides may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of  
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
 CC or lymphoma) and to stimulate the proliferation of T cells.  
 XX  
 SQ Sequence 18 AA;

Query Match 100.0%; Score 88; DB 21; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AADHRQLQLSTSSCLOQL 18  
 Db 1 aadhrqlqlstssclql 18

RESULT 2  
 Y06011  
 ID Y06011 standard; Peptide; 10 AA.

XX Y06011;  
 AC  
 DT 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.  
 OS  
 PN W09918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Rosenberg SA, Wang RF;

DR WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

PS Example 10; Page 42; 89pp; English.

XX

CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 24) corresponds to amino acid residues  
 CC 143-152 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 SQ Sequence 10 AA;

Query Match 54.5%; Score 48; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RQLQLSISSC 14  
 Db 1 rqlqlsissc 10

RESULT 3  
 Y06016  
 ID Y06016 standard; Peptide; 10 AA.

XX AC  
 AC Y06016;

XX DT 16-AUG-1999 (first entry)

XX DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

OS W09918206-A2.

XX PN

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenberg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

PS Example 10; Page 42; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 29) corresponds to amino acid residues  
 CC 144-153 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune

CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX Sequence 10 AA;

Query Match 53.4%; Score 47; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.05;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QLQLSISSCL 15  
 DB 1 qlqlsisscl 10  
 |||||

RESULT 4  
 Y06035  
 ID Y06035 standard; Peptide; 9 AA.

AC Y06035;  
 DT 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

OS Homo sapiens.

XX W09918206-A2.

PN 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

PF 08-OCT-1997; 97US-0061428.

PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA Rosenberg SA, Wang RF;

PI WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 18) corresponds to amino acid residues  
 CC 148-156 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,

CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX Sequence 9 AA;

Query Match 48.9%; Score 43; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SISSCLQQL 18  
 DB 1 sissclqql 9  
 |||||

RESULT 5  
 Y79748  
 ID Y79748 standard; Peptide; 9 AA.

AC Y79748;

XX 10-MAY-2000 (first entry)

XX NY-ESO-1 derived peptide #4.

XX Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;  
 KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;  
 KW cytostatic; melanoma; synovial sarcoma.

XX Homo sapiens.

XX W0200000824-A1.

XX 06-JAN-2000.

XX 25-JUN-1999; 99WO-US14493.

XX 26-JUN-1998; 98US-0105839.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;

PI Chen Y, Gure A, Old LJ;

XX WPI; 2000-170933/15.

XX Determining the possible presence of breast, endometrial, colorectal,  
 PT lung, bladder or head-neck cancer -

XX Example 13; Page 26; 40pp; English.

XX A method has been developed for determining the possible presence of a  
 CC cancer, which is not melanoma or synovial sarcoma. The method comprises  
 CC assaying a sample taken from the subject to determine the expression of  
 CC an SSX gene, and determining the expression of SSX1 gene indicates possible  
 CC presence of cancer. Expression of SSX1 gene indicates possible  
 CC presence of breast, endometrial, colorectal, lung, bladder or head-neck  
 CC cancer. These cancers are also detected by SSX2 and SSX4 gene expression.  
 CC SSX2 gene expression additionally indicates possible presence of  
 CC lymphoma, renal cell cancer, glioma and prostate cancer. Expression of  
 CC SSX4 gene also indicates possible presence of ovarian or stomach cancer.  
 CC SSX5 gene expression indicates the same cancers as SSX1, except breast  
 CC cancer. Determining expression of SSX gene can be used to monitor  
 CC progress of melanoma or synovial sarcoma, which is not cancer. The  
 CC SSX-derived peptide complex stimulates proliferation of cytolytic T  
 CC cells. This is useful for treating cancer, especially melanoma. Y78464  
 CC to Y78468 represent specifically claimed HLA binding peptides for use in  
 CC the method of the invention. Z88452 to Z88465 represent PCR primers used  
 CC in the isolation of SSX genes in the exemplification of the present  
 CC invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides  
 CC derived from SSX proteins or NY-ESO-1, which are used in the

CC exemplification of the present invention.

XX Sequence 9 AA;

Query Match 48.9%; Score 43; DB 21; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SSSCLOQL 18

Db 1 sssclql 9

RESULT 6

ID Y06037 standard; Peptide; 9 AA.

XX AC Y06037;

DT 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

OS WO9918206-A2.

PN 15-APR-1999.

PD 21-SEP-1998; 98WO-US19609.

PF 08-OCT-1997; 97US-0061428.

PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA Rosenberg SA, Wang RF;

PI WPI; 1999-277270/23.

PS Cancer antigen NY ESO1/CAG-3

PT Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 20) corresponds to amino acid residues  
 CC 145-133 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 9 AA;

Query Match 47.7%; Score 42; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LQLSISSCL 15

Db 1 lqlsisscl 9

RESULT 7

ID Y06042 standard; Peptide; 9 AA.

XX AC Y06042;

XX 16-AUG-1999 (first entry)

DT Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

OS WO9918206-A2.

PN 15-APR-1999.

PD 21-SEP-1998; 98WO-US19609.

PF 08-OCT-1997; 97US-0061428.

PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA Rosenberg SA, Wang RF;

PI WPI; 1999-277270/23.

PS Cancer antigen NY ESO1/CAG-3

PT Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 25) corresponds to amino acid residues  
 CC 143-151 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 9 AA;

Query Match 44.3%; Score 39; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RQLQLSISS 13

Db 1 rqlqlsiss 9

```
RESULT 8
Y79754
ID Y79754 standard; Peptide; 9 AA.
XX AC
XX Y79754;
XX AC
XX Y79754;
DT 10-MAY-2000 (first entry)
XX DE
XX NY-ESO-1 derived peptide #10.
XX DE
XX Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;
KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;
KW Cytostatic; melanoma; synovial sarcoma.
XX DE
XX Homo sapiens.
XX OS
XX WO200000824-A1.
XX PN
XX 06-JAN-2000.
XX PD
XX 25-JUN-1999; 99WO-US14493.
XX PF
XX 26-JUN-1998; 98US-0105839.
XX PR
XX (LUDW-) LUDWIG INST CANCER RES.
XX PA
XX Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;
PI Chen Y, Gure A, Old LJ;
PI PI
XX WPI; 2000-170933/15.
XX DR
XX Determining the possible presence of breast, endometrial, colorectal,
XX PT lung, bladder or head-neck cancer -
XX PT
XX Example 13; Page 26; 40pp; English.
XX PS
XX A method has been developed for determining the possible presence of a
XX CC cancer, which is not melanoma or synovial sarcoma. The method comprises
XX CC assaying a sample taken from the subject to determine the expression of
XX CC an SSX gene, and determining the expression as a determination of the
XX CC possible presence of cancer. Expression of SSX1 gene indicates possible
XX CC presence of breast, endometrial, colorectal, lung, bladder or head-neck
XX CC cancer. These cancers are also detected by SSX2 and SSX4 gene expression.
XX CC SSX2 gene expression additionally indicates possible presence of
XX CC lymphoma, renal cell cancer, glioma and prostate cancer. Expression of
XX CC SSX4 gene also indicates possible presence of ovarian or stomach cancer.
XX CC SSX5 gene expression indicates the same cancers as SSX1, except breast
XX CC cancer. Determining expression of SSX gene can be used to monitor
XX CC progress of melanoma or synovial sarcoma, which is not cancer. The
XX CC SSX-derived peptide complex stimulates proliferation of cytolytic T
XX CC cells. This is useful for treating cancer, especially melanoma. Y78464
XX CC to Y78468 represent specifically claimed HLA binding peptides for use in
XX CC the method of the invention. 288452 to 288465 represent PCR primers used
XX CC in the isolation of SSX genes in the exemplification of the present
XX CC invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides
XX CC derived from SSX proteins or NY-ESO-1, which are used in the
XX CC exemplification of the present invention.
XX CC
XX Sequence 9 AA;
XX SQ
```

```
Query Match 40.9%; Score 36; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADHROL 7
Db 3 aadhrol 9
|||||
3 aadhrol 9

RESULT 9
XX SQ
```

```
Query Match 40.9%; Score 36; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADHROL 7
Db 3 aadhrol 9
|||||
3 aadhrol 9

RESULT 9
XX SQ
```

```
Query Match 39.8%; Score 35; DB 13; Length 21;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AADHROLQSLIS 12
Db 4 apdhqelndvs 15
|||||
4 apdhqelndvs 15

RESULT 10
Y25818
ID Y25818 standard; Protein; 23 AA.
XX AC
XX Y25818;
XX AC
XX 04-OCT-1999 (first entry)
XX DT
XX DE
XX DE
XX Human secreted protein fragment encoded from gene 34.
XX KW
XX Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;
XX KW neurodegenerative disorder; developmental abnormality; blood disorder;
XX KW fetal deficiency; blood disorder; leukemia; immune system; inflammation;
XX KW autoimmune disease; hepatic disease; renal disease; allergy; restenosis;
XX KW ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
XX KW cardiovascular disorder; wound healing; stroke; arthritis; obesity;
XX KW asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
XX KW metabolic disorder.
XX KW
XX Homo sapiens.
XX XX
```

PN W09938881-A1.  
 XX  
 FD 05-AUG-1999.  
 XX  
 PF 27-JAN-1999; 99WO-US01621.  
 XX  
 PR 30-JAN-1998; 98US-00731170.  
 PR 30-JAN-1998; 98US-00731159.  
 PR 30-JAN-1998; 98US-00731160.  
 PR 30-JAN-1998; 98US-00731161.  
 PR 30-JAN-1998; 98US-00731162.  
 PR 30-JAN-1998; 98US-00731164.  
 PR 30-JAN-1998; 98US-00731165.  
 PR 30-JAN-1998; 98US-00731167.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Carter KC, Endress GA, Feng P, Ferrie AM, Florence C;  
 PI Florence KA, Janat F, Ni J, Rosen CA, Ruben SM;  
 PI Soppet DR, Young P, Yu G;  
 XX  
 DR WPI; 1999-469315/39.  
 DR N-PSDB; 200443.  
 XX  
 XX New isolated human genes and the secreted polypeptides they encode  
 PT useful in, e.g. treatment of Alzheimer's  
 PT  
 XX Disclosure; Page 357; 393pp; English.  
 PS  
 XX This invention describes novel human genes (see 200410-200477) and the  
 CC secreted proteins (see Y25711-Y25778) and fragments (see Y25779-Y25907)  
 CC they encode. The polynucleotides and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the 67  
 CC polynucleotides of the invention, based on which tissues they are most  
 CC highly expressed in, and include developing products for the diagnosis or  
 CC treatment of cancer, tumours, neurodegenerative disorders, developmental  
 CC abnormalities and fetal deficiencies, blood disorders, leukemias,  
 CC diseases of the immune system, autoimmune diseases, hepatic and renal  
 CC diseases, inflammation, allergies, ischaemic shock, Alzheimer's and  
 CC cognitive disorders, schizophrenia, restenosis, cardiovascular disorders,  
 CC wound healing, stroke, arthritis, obesity, asthma, sepsis, acne,  
 CC psoriasis, transplant rejection, metabolic disorders, infections and  
 CC AIDS. The polypeptides are also useful for identifying their binding  
 CC partners.  
 XX  
 SQ Sequence 23 AA;  
  
 Query Match 37.5%; Score 33; DB 20; Length 23;  
 Best Local Similarity 72.7%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 5 RQLQLSISSCL 15  
 Db 12 rplhlgsscl 22  
  
 RESULT 11  
 Y52440  
 ID Y52440 standard; Protein; 18 AA.  
 XX  
 AC Y52440;  
 XX  
 DT 15-FEB-2000 (first entry)  
 XX  
 DE Human tumour antigen NY-ESO-1 peptide #13.  
 XX  
 KW Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
 KW T-cell; helper; stimulation; proliferation; treatment;

KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
 KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
 KW lymphoma.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN W09953938-A1.  
 XX  
 PD 28-OCT-1999.  
 XX  
 PF 24-MAR-1999; 99WO-US06875.  
 XX  
 PR 17-APR-1998; 98US-0062422.  
 PR 02-OCT-1998; 98US-0165546.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
 PI Gure A, Ritter G;  
 XX  
 DR WPI; 2000-038483/03.  
 XX  
 PT Novel peptides which bind to MHC class I and MHC class II molecules,  
 PT useful for therapeutic and diagnostic purposes  
 PS  
 PS Claim 4; Page 22; 49pp; English.  
 XX  
 CC Peptides #8-13 (Y52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels  
 CC in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and  
 CC tumour cell lines with some degree of frequency - these included  
 CC melanoma specimens and cell lines, and breast and bladder cancer  
 CC specimens, with expression in other tumour types being sporadic.  
 CC These NY-ESO-1-derived peptides may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of  
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
 CC or lymphoma) and to stimulate the proliferation of T cells.  
 XX  
 SQ Sequence 18 AA;  
  
 Query Match 36.4%; Score 32; DB 21; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 AADHRQ 6  
 Db 13 aadhrq 18  
  
 RESULT 12  
 Y45398  
 ID Y45398 standard; Peptide; 13 AA.  
 XX  
 AC Y45398;  
 XX  
 DT 01-DEC-1999 (first entry)  
 XX  
 DE Immunogenic peptide having a human leukocyte antigen binding motif #9.  
 XX  
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.  
 XX

OS Synthetic.  
 XX Homo sapiens.  
 PN WO9945954-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 13-MAR-1998; 98WO-US05039.  
 XX  
 PR 13-MAR-1998; 98WO-US05039.  
 XX  
 XX (EPIM-) EPIMUNE INC.  
 XX  
 PI Sette A, Kubo RT, Sidney J, Cellis E, Grey HM, Southwood S;  
 XX WPI; 1999-551214/46.  
 DR  
 XX  
 PT New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases  
 XX  
 PS Claim 1; Page 27; 150pp; English.  
 XX  
 CC Y45390 to Y48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.  
 XX  
 SQ Sequence 13 AA;

Query Match 33.0%; Score 29; DB 20; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 STSSCLQQ 17  
 | | | | |  
 Db 1 stssclhq 8

RESULT 13  
 R84751  
 ID R84751 standard; peptide; 18 AA.  
 XX  
 AC R84751;  
 XX  
 XX 21-MAY-1996 (first entry)  
 DT  
 DE  
 DE GST-phospholipase C-gamma-SH3 fusion protein fragment #1.  
 XX  
 KW Ribonucleoprotein; DYN domain; signal transduction; SH3 domain; therapy;  
 KW neurodegenerative disorder; neuroproliferative disorder; cancer; dynamin.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 10

/note= "any amino acid"

WO9526983-A2.  
 12-OCT-1995.  
 30-MAR-1995; 95WO-US03945.  
 15-AUG-1994; 94US-0291591.  
 31-MAR-1994; 94US-0221642.  
 31-MAY-1994; 94US-0251691.  
 (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 Hobert O, Jallal B, Kostka G, Obermeier A, Ullrich A;  
 WPI; 1995-366151/47.  
 Treatment of a disease or condition characterised by abnormality in  
 a signal transduction pathway - by disrupting or promoting the  
 interaction in vivo  
 Example 1; Page 70; 100pp; English.  
 R84751 and R84752 represent fragments of a GST-phospholipase C-gamma-SH3  
 fusion protein. This sequence was screened to see if it could be used  
 for the treatment of a disease where there is an abnormality in a signal  
 transduction pathway. The screening method looks for a disruption (or  
 promotion) of the interaction between this sequence and an SH3 domain.  
 The sequence can also be used for screening non-haematopoietic cells for  
 a protein with an SH3 domain. The methods can be used for screening,  
 diagnosing, and treating diseases, such as neurodegenerative or  
 neuroproliferative disorders, or cancers.  
 Sequence 18 AA;

Query Match 33.0%; Score 29; DB 16; Length 18;  
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 DHRQLQLS 10  
 | : | | | : |  
 Db 2 dyrqlqla 9

RESULT 14  
 W96889  
 ID W96889 standard; peptide; 18 AA.  
 XX  
 AC W96889;  
 XX  
 DT 22-APR-1999 (first entry)  
 XX  
 DE ApoB-100 nuclear localisation signal sequence, residues 2228-2245.  
 XX  
 KW Human apolipoprotein B-100; apoB-100; very-low density lipoprotein; VLDL;  
 KW apolipoprotein; binding; in vivo transport; nucleic acid; binding domain;  
 KW nuclear localisation sequence; gene therapy; cancer; cystic fibrosis;  
 KW non-small cell lung carcinoma; diabetes; arteriosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9856938-A1.  
 XX  
 PD 17-DEC-1998.  
 XX  
 PF 10-JUN-1998; 98WO-US11927.  
 XX  
 PR 14-MAY-1998; 98US-0079030.  
 PR 13-JUN-1997; 97US-0874807.  
 XX  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.

XX Guevara JG, Hoogveen RC, Moore JP;  
XX WPI; 1999-070331/06.  
XX Composition comprising nucleic acid bound to LDL or VLDL lipoprotein  
PT - used for delivering nucleic acid to cells for gene therapy and  
PT antisense treatment  
XX  
PS Claim 19; Fig 13D; 293pp; English.  
XX  
CC W96878-97 represent nuclear localisation signal sequence derived from  
CC human apolipoprotein B-100 (apoB-100). ApoB-100 is a major apoprotein  
CC component of very-low density lipoproteins (VLDL), intermediate density  
CC lipoprotein (IDL), low density lipoproteins (LDL) and lipoprotein a.  
CC The present sequence can be used in the composition of the invention.  
CC The specification describes a composition that comprises LDL and  
CC apolipoproteins for the binding and in vivo transport of nucleic acids.  
CC The composition is used to deliver nucleic acids to eukaryotic cells,  
CC in vivo or in vitro, for expressing a therapeutic polypeptide or  
CC antisense molecule (or ribozyme). Specifically they are used for gene  
CC therapy of cancers (particularly non-small cell lung carcinoma),  
CC diabetes, cystic fibrosis and arteriosclerosis.  
XX  
SQ Sequence 18 AA;  
  
Query Match 33.0%; Score 29; DB 20; Length 18;  
Best Local Similarity 46.2%; Pred. No. 1.3e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
QY 6 QLQLSISSCLOQL 18  
Db 3 qiriqlqeklql 15  
  
RESULT 15  
W11227  
ID W11227 standard; Peptide; 20 AA.  
XX  
AC W11227;  
XX  
DT 17-JUN-1997 (first entry)  
XX  
DE Modified rat obese (ob) protein fragment (1).  
XX  
DE Modified; rat; obese; ob; fragment; antigen; antibody;  
KW detection; assay; purification.  
XX  
OS Rattus rattus.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT Peptide /note= "acylated"  
FT 1..20  
FT /note= "residues 148-167 of rat ob protein"  
XX  
XX WO9631526-A1.  
XX  
XX 10-OCT-1996.  
XX  
XX 05-APR-1996; 96WO-US04909.  
XX  
XX 06-APR-1995; 95US-0419598.  
XX  
XX (AMYL-) AMYLIN PHARM INC.  
XX  
PI Albrandt KA, Beeley N, Beidler DE, Chun M, Janes SM;  
PI Park DM, Phelps JL, Prickett KS, Rink TJ, Sierzega MB;  
XX WPI; 1996-464965/46.  
XX  
PT Modified obese (ob) gene prods. - comprising ob dimers, ob dimer

PT fusion proteins and ob monomer fusion proteins, useful, e.g. for  
PT treating obesity or diabetes  
XX  
PS Example 8; Page -; 130pp; English.  
XX  
CC The present sequence is a modified rat obese (ob) protein  
CC fragment based on residues 148-167, which can be used as an antigen  
CC in the preparation of antibodies (Ab) against the rat ob  
CC protein. The Ab can be used for detection, assays and  
CC purifications.  
XX  
SQ Sequence 20 AA;  
  
Query Match 33.0%; Score 29; DB 17; Length 20;  
Best Local Similarity 53.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 6 QLQLSISSCLOQL 18  
Db 2 riqqalqdlqlql 14

Search completed: May 8, 2001, 15:07:07  
Job time: 223 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:08:18 ; Search time 68.24. Seconds  
(without alignments)  
18.127 Million cell updates/sec

Title: US-09-165-546A-8

Perfect score: 88

Sequence: 1 AADHRQLQLSISSCLQL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 4929

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_67:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	36.4	23	2 S72535	probable acr-2 reg
2	30	34.1	24	4 S09363	hypothetical MTCOL
3	28	31.8	23	2 S31210	collagen alpha 2(I
4	25	28.4	20	2 A35857	arginine--CRNA lig
5	24	27.3	16	2 S34444	blaz protein - Sta
6	24	27.3	17	2 S05917	chorion class A pr
7	24	27.3	23	2 S60570	homeodomain protei
8	24	27.3	24	2 S01808	hemoglobin AII - t
9	23	26.1	14	2 I49514	B144 protein A - m
10	23	26.1	15	2 S28873	lipid transfer pro
11	23	26.1	17	2 B23219	high-cysteine chor
12	23	26.1	17	2 A29834	trp leader peptide
13	23	26.1	21	2 S46550	actin-related prot
14	23	26.1	22	2 B41833	40K iron-repressed
15	23	26.1	24	2 T17073	NADH dehydrogenase
16	22	25.0	14	2 I54284	Cl-inhibitor - hum
17	22	25.0	15	2 D48394	major fat-globule
18	22	25.0	17	2 S05913	chorion class A pr
19	22	25.0	20	2 A60489	vencombin A (EC 3.4
20	22	25.0	21	2 G64564	hypothetical prote
21	22	25.0	21	2 H64640	hypothetical prote
22	22	25.0	21	2 H71858	hypothetical prote
23	22	25.0	23	2 S60564	homeodomain protei
24	21	23.9	12	2 I57678	gene rPLP-A protei
25	21	23.9	14	2 PH1806	T cell receptor al
26	21	23.9	18	2 S63487	fibrinogenolytic p
27	21	23.9	18	2 S59490	translation elonga
28	21	23.9	19	2 A44854	L-2,4-diaminobuty
29	21	23.9	19	2 PQ0678	photosystem I 8.0K

30 21 23.9 20 2 B53875 creatine kinase (E  
31 21 23.9 20 2 S14161 probable dTDP-4-ke  
32 21 23.9 21 2 S68971 hypothetical prote  
33 21 23.9 21 2 A53630 anti-HIV antidioti  
34 21 23.9 22 2 A41833 40K iron-repressed  
35 21 23.9 22 2 F84018 hypothetical prote  
36 21 23.9 24 2 S30923 hypothetical prote  
37 21 23.9 24 2 A58999 psi-conotoxin PIII  
38 21 23.9 25 2 D47689 flagellar core pro  
39 20 22.7 6 2 S71349 beta-crystallin B2  
40 20 22.7 13 2 PT0305 ig heavy chain CRD  
41 20 22.7 14 2 PH1321 ig heavy chain DJ  
42 20 22.7 15 2 PT0096 pyruvate dehydroge  
43 20 22.7 17 2 S05923 chorion class A pr  
44 20 22.7 17 2 S7519 T cell receptor be  
45 20 22.7 18 2 A41877 LcrKc - Yersinia p

#### ALIGNMENTS

##### RESULT 1

S72535  
probable acr-2 regulatory leader protein - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 21-Jul-2000  
C;Accession: S72535  
R;Akiyama, M.; Nakashima, H.  
Biochim. Biophys. Acta 1307, 187-192, 1996  
A;Title: Molecular cloning of the acr-2 gene which controls acriflavine sensitivity i  
A;Reference number: S72535; MUID:96283814  
A;Accession: S72535  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-23 <AKT>  
A;Cross-references: EMBL:D45893; NID:gl754593; PIDN:BAA08306.1; PID:gl754594

Query Match 36.4%; Score 32; DB 2; Length 23;  
Best Local Similarity 50.0%; Pred. No. 29;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 RQLQLSISSC 14  
| : : | : |||  
DB 11 RMIRLEVSSC 20

##### RESULT 2

S09363  
hypothetical MTCOL/MTCYB mutant fusion protein - human mitochondrion (fragment)  
C;Species: mitochondrion Homo sapiens (man)  
C;Date: 21-Nov-1993 #sequence\_revision 14-Aug-1997 #text\_change 20-Apr-2000  
C;Accession: S09363  
R;Poulton, J.; Deadman, M.E.; Gardiner, R.M.  
Nucleic Acids Res. 17, 10223-10229, 1989  
A;Title: Tandem direct duplications of mitochondrial DNA in mitochondrial myopathy: a  
A;Reference number: S09363; MUID:90098864  
A;Accession: S09363  
A;Molecule type: DNA  
A;Residues: 1-24 <POU>  
C;Comment: This is the hypothetical translation of a sequence believed to result from  
C;Genetics:  
A;Gene: MTCOL/MTCYB  
A;Genome: mitochondrion  
A;Genetic code: SGC1  
C;Keywords: fusion protein; mitochondrion  
F;1-4/Region: cytochrome-c oxidase chain I  
F;5-24/Region: cytochrome b (+2 frame shifted)

Query Match 34.1%; Score 30; DB 4; Length 24;  
Best Local Similarity 44.4%; Pred. No. 69;  
Matches 8; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 3 DHRQLQ----LSISSCLQ 16  
||| | :|||||  
Db 6 DHFSTQKPEALSCLQ 23

## RESULT 3

S31210  
collagen alpha 2(I) chain precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 13-Aug-1999  
C:Accession: S31210  
R:Guenette, D.K.; Ritzenthaler, J.D.; Foley, J.; Jackson, J.D.; Smith, B.D.  
Biochem. J. 283, 699-703, 1992  
A:Title: DNA methylation inhibits transcription of procollagen alpha-2(I) promoters.  
A:Reference number: S31210; MUID:92272666  
A:Accession: S31210  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-23 <GUE>  
A:Cross-references: EMBL:X66209; NID:g55974; PIDN:CAA46960.1; PID:g55975  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix

Query Match 31.8%; Score 28; DB 2; Length 23;  
Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 7 IQLSISSCL 15  
| | :|||  
Db 11 LLLAVTSCL 19

## RESULT 4

A35857  
arginine-tRNA ligase (EC 6.1.1.19) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 12-Oct-1990 #sequence\_revision 12-Oct-1990 #text\_change 14-Nov-1997  
C:Accession: A35857  
R:Siivaram, P.; Deutscher, M.P.  
Proc. Natl. Acad. Sci. U.S.A. 87, 3665-3669, 1990  
A:Title: Existence of two forms of rat liver arginyl-tRNA synthetase suggests channeling  
A:Reference number: A35857; MUID:90251617  
A:Accession: A35857  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <SIV>  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 28.4%; Score 25; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 4.6e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 9 LSISSCLOQL 18  
:| | | |  
Db 2 ININSXLQEL 11

## RESULT 5

S34444  
blaz protein - Staphylococcus aureus plasmid pI258 (fragment)  
C:Species: Staphylococcus aureus  
C>Date: 22-Nov-1993 #sequence\_revision 09-Mar-1996 #text\_change 07-May-1999  
C:Accession: S34444  
R:Wang, P.Z.; Projan, S.J.; Novick, R.P.  
Nucleic Acids Res. 19, 4000, 1991  
A:Title: Nucleotide sequence of beta-lactamase regulatory genes from staphylococcal plasmid pI258  
A:Reference number: S34444; MUID:91319567  
A:Accession: S34444  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-16 <WAN>  
A:Cross-references: EMBL:M62650  
A:Experimental source: strain RN11  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199  
C:Genetics:  
A:Gene: blaz  
A:Genome: plasmid pI258

Query Match 27.3%; Score 24; DB 2; Length 16;  
Best Local Similarity 71.4%; Pred. No. 5.5e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 RQLQLSI 11  
:| | | |  
Db 6 RQLQLXI 12

## RESULT 6

S05917  
chorion class A protein L4 precursor - silkworm (fragment)  
C:Species: Bombyx mori (silkworm)  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Aug-1999  
C:Accession: S05917  
R:Speerel, N.A.; Nguyen, H.T.; Eickbush, T.H.; Kafatos, F.C.  
J. Mol. Biol. 209, 1-19, 1989  
A:Title: Gene evolution and regulation in the chorion complex of Bombyx mori. Hybrid  
A:Reference number: S05913; MUID:90040707  
A:Accession: S05917  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-17 <SPO>  
A:Cross-references: EMBL:X15560; NID:g5781; PIDN:CAA33571.1; PID:g5782  
A:Note: this sequence was submitted to the EMBL Data Library, Jun-1989  
C:Genetics:  
A:Map position: 2  
C:Superfamily: chorion class A protein pc292

Query Match 27.3%; Score 24; DB 2; Length 17;  
Best Local Similarity 45.5%; Pred. No. 5.8e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 IQLSISSCLQ 17  
| | :|||  
Db 7 LLLCVQACLIQ 17

## RESULT 7

S60570  
homeodomain protein hrox8 - California red abalone (fragment)  
C:Species: Haliotis rufescens (California red abalone)  
C>Date: 19-Mar-1997 #sequence\_revision 09-May-1997 #text\_change 15-Oct-1999  
C:Accession: S60570  
R:Degnan, B.M.; Morse, D.E.  
Mol. Marine Biol. Biotechnol. 2, 1-9, 1993  
A:Title: Identification of eight homeobox-containing transcripts expressed during lar  
A:Reference number: S60564; MUID:93372986  
A:Accession: S60570  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-23 <DEG>  
A:Cross-references: EMBL:X79377; NID:g495120; PIDN:CAA55922.1; PID:g495121  
C:Superfamily: unassigned homeobox proteins; homeobox homology  
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 27.3%; Score 24; DB 2; Length 23;  
Best Local Similarity 23.1%; Pred. No. 8e+02;  
Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 5 RQLQLSISSCLQ 17  
| : : : : | | :

Db 9 RRTEIAHTLCISE 21

RESULT 8

S01808  
hemoglobin AII - tube worm (Lamellibrachia 'sp.') (fragment)  
C:Species: Lamellibrachia sp.  
C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 04-Mar-2000  
C:Accession: S01808  
R:Suzuki, T.; Takagi, T.; Ohta, S.  
Biochem. J. 255, 541-545, 1988  
A:Title: N-terminal amino acid sequence of the deep-sea tube worm haemoglobin remarkably  
A:Reference number: S01807; MUID:89076216  
A:Accession: S01808  
A:Molecule type: protein  
A:Residues: 1-24 <SUZ>  
C:Superfamily: globin; globin homology  
C:Keywords: oxygen carrier

Query Match 27.3%; Score 24; DB 2; Length 24;

Best Local Similarity 57.1%; Pred. No. 8.3e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 DHRLQL 9

I : I I I

Db 9 DRREMQL 15

RESULT 9

I49514  
B144 protein A - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I49514  
R:Tsuze, I.; Shen, F.  
Immunogenetics 26, 378-380, 1987  
A:Title: A gene in the H-2S: H-2D interval of the major histocompatibility complex which  
A:Reference number: I49514; MUID:88031493  
A:Accession: I49514  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-14 <RES>  
A:Cross-references: GB:M18187; NID:g192097; PIDN:AAA37272.1; PID:g192098

Query Match 26.1%; Score 23; DB 2; Length 14;

Best Local Similarity 40.0%; Pred. No. 7.2e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 LSISSCLOQL 18

I : I I I

Db 2 LDLQACSQSL 11

RESULT 10

S28873  
lipid transfer protein Cw(41) - maize (fragment)  
C:Species: Zea mays (maize)  
C:Date: 19-Mar-1997 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998  
C:Accession: S28873  
R:Molina, A.; Segura, A.; Garcia-Olmedo, F.  
FEBS Lett. 316, 119-122, 1993  
A:Title: Lipid transfer proteins (nsLTPs) from barley and maize leaves are potent inhibi  
A:Reference number: S28871; MUID:93131027  
A:Accession: S28873  
A:Molecule type: protein  
A:Residues: 1-15 <MOL>  
A:Experimental source: etiolated leaf, cultivar INIA 1986  
C:Keywords: lipid transport

Query Match 26.1%; Score 23; DB 2; Length 15;

Best Local Similarity 40.0%; Pred. No. 7.7e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 QLQLSISSCL 15

I : I I I

Db 6 QVSXAIAPCL 15

RESULT 11

B23219  
high-cysteine chorion protein A 13 - silkworm (fragment)  
C:Species: Bombyx mori (silkworm)  
C:Date: 22-Jul-1987 #sequence\_revision 22-Jul-1987 #text\_change 21-Jul-2000  
C:Accession: B23219  
R:Rodakis, G.C.; Lecanidou, R.; Eickbush, T.H.  
J. Mol. Evol. 20, 265-273, 1984  
A:Title: Diversity in a chorion multigene family created by tandem duplications and a  
A:Reference number: A2960; MUID:85083111  
A:Accession: B23219  
A:Molecule type: DNA  
A:Residues: 1-17 <ROD>  
A:Cross-references: GB:X01068; NID:g5752; PIDN:CAB57790.1; PID:g6015486  
C:Genetics: 17/3  
A:Introns: 17/3  
C:Superfamily: chorion class A protein pc292

Query Match 26.1%; Score 23; DB 2; Length 17;

Best Local Similarity 45.5%; Pred. No. 8.8e+02;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 LQLSISSCLQ 17

I : I I I

Db 7 LLLCVQCCLIQ 17

RESULT 12

A29834  
trp leader peptide - Corynebacterium glutamicum  
C:Species: Corynebacterium glutamicum  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 24-Sep-1999  
C:Accession: A29834; A24723; A29458; S13087; A48967  
R:Matsui, K.; Miwa, K.; Sano, K.  
J. Bacteriol. 169, 5330-5332, 1987  
A:Title: Two single-base-pair substitutions causing desensitization to tryptophan fee  
ntum  
A:Reference number: A29834; MUID:88032866  
A:Contents: B. lactofermentum  
A:Accession: A29834  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-17 <MAT>  
A:Cross-references: GB:M17892; NID:g144101; PIDN:AAB59110.1; PID:g1129101  
R:Matsui, K.; Sano, K.; Ohtsubo, E.  
Nucleic Acids Res. 14, 10113-10114, 1986  
A:Title: Complete nucleotide and deduced amino acid sequences of the Brevibacterium 1  
A:Reference number: A93606; MUID:87117512  
A:Contents: B. lactofermentum  
A:Accession: A24723  
A:Molecule type: DNA  
A:Residues: 1-17 <MA2>  
A:Cross-references: GB:X04960; NID:g39591; PIDN:CAA28622.1; PID:g580785  
R:Sano, K.; Matsui, K.  
Gene 53, 191-200, 1987  
A:Title: Structure and function of the trp operon control regions of Brevibacterium 1  
A:Reference number: A91575; MUID:87277409  
A:Contents: B. lactofermentum  
A:Accession: A29458  
A:Molecule type: DNA  
A:Residues: 1-17 <SAN>  
R:Heery, D.M.; Dunican, L.K.  
Nucleic Acids Res. 18, 7138, 1990  
A:Title: Nucleotide sequence of the Corynebacterium glutamicum trpE gene.

A;Reference number: S13087; MUID:91088299

A;Accession: S13087

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-17 <HEE>

A;Cross-references: EMBL:X55994; NID:g40521; PIDN:CAA39466.1; PID:g580992

R;Herry, D.M.; Dunican, L.K.

Appl. Environ. Microbiol. 59, 791-799, 1993

A;Title: Cloning of the trp gene cluster from a tryptophan-hyperproducing strain of Cory

A;Reference number: A48967; MUID:93243735

A;Accession: A48967

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-13 <HER>

A;Cross-references: GB:S59299; NID:g299877; PIDN:AAB26395.1; PID:g299878

A;Experimental source: ATCC 21850

A;Note: sequence extracted from NCBI backbone (NCBIN:130455, NCBIP:130456)

C;Genetics:

A;Gene: trpL

A;Start codon: GTG

C;Superfamily: unassigned leader peptides

Query Match 26.1%; Score 23; DB 2; Length 17;

Best Local Similarity 66.7%; Pred. No. 8.8e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 12 SSCLOQ 17

Db 3 NSCLSQ 8

RESULT 13

S46550

actin-related protein - chicken (fragments)

C;Species: Gallus gallus (chicken)

C;Date: 12-Sep-1995 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999

A;Accession: S46550

R;Lees-Miller, J.P.; Helfman, D.M.; Schroer, T.A.

Nature 359, 244-246, 1992

A;Title: A vertebrate actin-related protein is a component of a multisubunit complex inv

A;Reference number: S29089; MUID:92408781

A;Accession: S46550

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-21 <LEE>

Query Match 26.1%; Score 23; DB 2; Length 21;

Best Local Similarity 50.0%; Pred. No. 1.1e+03;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AADHRLQ 8

Db 9 AEHRKAQ 16

RESULT 14

B41833

40K iron-repressed periplasmic protein - Haemophilus influenzae (strain PAK12085) (fragm

C;Species: Haemophilus influenzae

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change: 11-Jan-2000

A;Accession: B41833

R;Harkness, R.E.; Chong, P.; Klein, M.H.

J. Bacteriol. 174, 2425-2430, 1992

A;Title: Identification of two iron-repressed periplasmic proteins in Haemophilus influe

A;Reference number: A41833; MUID:92210482

A;Accession: B41833

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-22 <HAR>

C;Superfamily: sfuA protein

Query Match 26.1%; Score 23; DB 2; Length 22;

Best Local Similarity 14.3%; Pred. No. 1.2e+03;

Matches 2; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 4 HRQLSLSSCLQQ 17

Db 9 HRAAAVAVAKAFEQ 22

RESULT 15

T17073

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - Chamaeleo fischeri mitochondri

C;Species: mitochondrion Chamaeleo fischeri

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000

A;Accession: T17073

R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Papenfuss, T.J.

J. Mol. Evol. 44, 660-674, 1997

A;Title: Evolutionary shifts in three major structural features of the mitochondrial

A;Reference number: Z18674; MUID:97315309

A;Accession: T17073

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-24 <MAC>

A;Cross-references: EMBL:U82688; NID:g3603112; PID:g3603113; PIDN:AAC62273.1

C;Genetics:

A;Genome: mitochondrion

A;Note: ND1

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 1

C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylato

Query Match 26.1%; Score 23; DB 2; Length 24;

Best Local Similarity 44.4%; Pred. No. 1.3e+03;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LOLSISL 15

Db 3 LPLTLAMCL 11

Search completed: May 8, 2001, 15:08:19

Job time: 290 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:05 ; Search time 40.07 Seconds  
(without alignments)  
15.388 Million cell updates/sec

Title: US-09-165-546A-8

Perfect score: 88

Sequence: 1 AADHRQLQLSISSCLQQL 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	28.4	20	1	SYR_RAT
2	23	26.1	17	1	CHH3_BOMMO
3	23	26.1	17	1	LPW_BELA
4	23	26.1	19	1	PYRB_PSEFL
5	22	25.0	20	1	ANCR_AKGBI
6	21	23.9	21	1	RP71_STRSQ
7	21	23.9	25	1	FLB3_TREHY
8	20	22.7	9	1	OXYT_RAJCL
9	20	22.7	17	1	PROP_SALTY
10	20	22.7	22	1	APC1_MACEA
11	20	22.7	22	1	CH60_BOVIN
12	20	22.7	24	1	ODPB_BOVIN
13	19	21.6	14	1	RECJ_SALTY
14	19	21.6	17	1	TPIS_PINPS
15	19	21.6	18	1	FIXA_RHILE
16	19	21.6	24	1	CH60_HELVI
17	19	21.6	24	1	LPTN_ECOLI
18	18	20.5	9	1	ISOT_CYPCA
19	18	20.5	10	1	FARP_WYTED
20	18	20.5	10	1	HTF_HELZE
21	18	20.5	13	1	HPA1_RANES
22	18	20.5	15	1	48KD_BACCE
23	18	20.5	16	1	BATI_EUBSP
24	18	20.5	16	1	YMOR_PSEPU
25	18	20.5	19	1	PHLC_STAIN
26	18	20.5	20	1	CPA7_PAPSP
27	18	20.5	20	1	NEUA_RAT
28	18	20.5	20	1	YOAH_KLEAE
29	18	20.5	22	1	MOT1_CANFA
30	18	20.5	24	1	VPB_BPD10
31	18	20.5	25	1	LVC_ASTRU
32	17	19.3	9	1	OXYT_BUFRE
33	17	19.3	12	1	UH03_RAT

RESULT	1	17	19.3	15	1	ALLS_MANSE	P42559	manduca sex
SYR_RAT	17	19.3	17	1	1	ERFQ_THEAQ	Q01697	thermus aqu
ID	AC	P40329;	AC	P40329;	STANDARD;	PRT;	20	AA.
DT	01-FEB-1995	(Rel. 31, Created)	DT	01-FEB-1995	(Rel. 31, Last sequence update)			
DT	01-FEB-1995	(Rel. 31, Last sequence update)	DT	01-FEB-1995	(Rel. 31, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)	DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	ARGINYL-TRNA SYNTHETASE (EC 6.1.1.19) (ARGININE--TRNA LIGASE) (ARGRS)		DE	ARGINYL-TRNA SYNTHETASE (EC 6.1.1.19) (ARGININE--TRNA LIGASE) (ARGRS)				
DE	(FRAGMENT)		DE	(FRAGMENT)				
GN	RARS.		GN	RARS.				
OS	Rattus norvegicus (Rat).		OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.		OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
ON	NCBI_TaxID=10116;		ON	NCBI_TaxID=10116;				
RN	[1]		RN	[1]				
RP	SEQUENCE.		RP	SEQUENCE.				
RX	MEDLINE=90251617; Pubmed=2187187;		RX	MEDLINE=90251617; Pubmed=2187187;				
RA	Sivaram P., Deutscher M.P.;		RA	Sivaram P., Deutscher M.P.;				
RT	"Existence of two forms of rat liver arginyl-trna synthetase suggests		RT	"Existence of two forms of rat liver arginyl-trna synthetase suggests				
RT	channeling of aminoacyl-trna for protein synthesis.";		RT	channeling of aminoacyl-trna for protein synthesis.";				
CC	Proc. Natl. Acad. Sci. U.S.A. 87:3665-3669(1990).		CC	Proc. Natl. Acad. Sci. U.S.A. 87:3665-3669(1990).				
CC	-1- CATALYTIC ACTIVITY: ATP + L-ARGININE + TRNA(ARG) = AMP +		CC	-1- CATALYTIC ACTIVITY: ATP + L-ARGININE + TRNA(ARG) = AMP +				
CC	PYROPHOSPHATE + L-ARGINYL-TRNA(ARG)		CC	PYROPHOSPHATE + L-ARGINYL-TRNA(ARG)				
CC	-1- SUBUNIT: MONOMER; ALSO PART OF A MULTISUBUNIT COMPLEX THAT GROUPS		CC	-1- SUBUNIT: MONOMER; ALSO PART OF A MULTISUBUNIT COMPLEX THAT GROUPS				
CC	TRNA LIGASES FOR ARG, ASP, GLU, GLN, ILE, LEU, LYS, MET AND PRO.		CC	TRNA LIGASES FOR ARG, ASP, GLU, GLN, ILE, LEU, LYS, MET AND PRO.				
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.		CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC.				
DR	PIR; A35857; A35857.		DR	PIR; A35857; A35857.				
DR	InterPro: IPR001412;		DR	InterPro: IPR001412;				
DR	PROSITE; PS00178; AA_TRNA_LIGASE_I; PARTIAL.		DR	PROSITE; PS00178; AA_TRNA_LIGASE_I; PARTIAL.				
KW	Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.		KW	Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.				
FT	NON_TER 20 20		FT	NON_TER 20 20				
SQ	SEQUENCE 20 AA; 2217 MW; D3DDA1847A5EB46B CRC64;		SQ	SEQUENCE 20 AA; 2217 MW; D3DDA1847A5EB46B CRC64;				

Query Match 28.4%; Score 25; DB 1; Length 20;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 LSISLQQL 18  
DB 2 ININSIQEL 11

RESULT 2

CHH3_BOMMO	STANDARD;	PRT;	17	AA.
ID	CHH3_BOMMO			
AC	P20729;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	01-FEB-1991 (Rel. 17, Last annotation update)			
DE	CHORION CLASS HIGH-CYSTEINE HCA PROTEIN 13 PRECURSOR (HC-A.13)			
DE	(FRAGMENT)			
OS	Bombyx mori (Silk moth).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;			

P42559 manduca sex  
Q01697 thermus aqu  
P82536 spinacia ol  
P15199 rabies viru  
P08497 bacillus su  
P04358 agarius bi  
P02807 neurospora  
P13066 desulfovibr  
P12381 klebsiella  
P81018 oncorhynchu  
P81022 polyodon sp  
P41032 salmonella

OC Bombycoidea: Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85083111; PubMed=6439880;  
RA Rodakis G.C., Lecanidou R., Eickbush T.H.;  
RT "Diversity in a chorion multigene family created by tandem  
RT duplications and a putative gene-conversion event.";  
RL J. Mol. Evol. 20:265-273(1984).  
CC -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGG SHELL OF THE  
CC SILK MOTH.  
CC -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH  
CC BELONG CLASSES A, CA AND HCA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X01068; CAB57790.1; -;  
DR PIR; B23219; B23219.  
KW Eggshell; Chorion; Repeat; Multigene family; Signal.  
FT SIGNAL 1  
FT NON\_TER 17  
SQ SEQUENCE 17 AA: 1913 MW; 5E634508C5355C9C CRC64;  
  
Query Match 26.1%; Score 23; DB 1; Length 17;  
Best Local Similarity 45.5%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
QY 7 LQLSSICLQ 17  
Db | | | | | | | | | |  
7 LLLCVQGLIQ 17  
  
RESULT 3  
LPW\_BRELA STANDARD; PRT; 17 AA.  
AC P06556;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE TRP OPERON LEADER PEPTIDE.  
GN TRPL  
OS Brevibacterium lactofermentum, and  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae.  
OX NCBI\_TaxID=92707, 1718;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=B.lactofermentum;  
RX MEDLINE=87117512; PubMed=3808947;  
RA Matsui K., Sano K., Ohtsubo E.;  
RT "Complete nucleotide and deduced amino acid sequences of the  
RT Brevibacterium lactofermentum tryptophan operon.";  
RL Nucleic Acids Res. 14:10113-10114(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=B.lactofermentum;  
RX MEDLINE=88032866; PubMed=3667535;  
RA Matsui K., Miwa K., Sano K.;  
RT "Two single-base-pair substitutions causing desensitization to  
RT tryptophan feedback inhibition of anthranilate synthase and enhanced  
RT expression of tryptophan genes of Brevibacterium lactofermentum.";  
RL J. Bacteriol. 169:5330-5332(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=B.lactofermentum;

RX MEDLINE=87277409; PubMed=3609747;  
RA Sano K., Matsui K.;  
RT "Structure and function of the trp operon control regions of  
RT Brevibacterium lactofermentum, a glutamic-acid-producing bacterium.";  
RL Gene 53:191-200(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=C.glutamicum; STRAIN=ATCC 13059 / AS019;  
RX MEDLINE=91088299; PubMed=2263476;  
RA Heery D.M., Dunican L.K.;  
RT "Nucleotide sequence of the Corynebacterium glutamicum trpE gene.";  
RL Nucleic Acids Res. 18:7138-7138(1990).  
CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS  
CC OF TRYPTOPHAN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X04960; CAA28622.1; -;  
DR EMBL; M17892; AAB59110.1; -;  
DR EMBL; M16663; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X55994; CAA39466.1; -;  
DR PIR; A29458; A29458.  
DR PIR; A29834; A29834.  
DR PIR; A24723; A24723.  
DR PIR; S13087; S13087.  
KW Tryptophan biosynthesis; Leader peptide.  
SQ SEQUENCE 17 AA; 2112 MW; 74C7E7924DAAE56B CRC64;  
  
Query Match 26.1%; Score 23; DB 1; Length 17;  
Best Local Similarity 56.7%; Pred. No. 3.3e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 12 SSCLOQ 17  
Db | | | | | | | | | |  
3 NSCLSQ 8  
  
RESULT 4  
PYRB\_PSEFL STANDARD; PRT; 19 AA.  
AC P56585;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ASPARTATE CARBAMOYLTRANSFERASE (EC 2.1.3.2) (ASPARTATE  
DE TRANSCARBAMYLASE) (ATCASE) (FRAGMENT).  
GN PYRB.  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=ATCC 13525;  
RX MEDLINE=94052183; PubMed=8234318;  
RA Bergh S.T., Evans D.R.;  
RT "Subunit structure of a class A aspartate transcarbamoylase from  
RT Pseudomonas fluorescens.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:9818-9822(1993).  
CC -!- CATALYTIC ACTIVITY: CARBAMOYL-PHOSPHATE + ASPARTATE =  
CC ORTHOPHOSPHATE + N-CARBAMOYLASPARTATE.  
CC -!- PATHWAY: SECOND STEP IN PYRIMIDINE BIOSYNTHESIS  
CC -!- SUBUNIT: HETERODIMER OF 6 ACTIVE PYRB SUBUNITS AND 6 NON-  
CC CATALYTIC PYRB SUBUNITS.  
CC -!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.  
DR InterPro; IPR002029; -;

DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; PARTIAL.  
KW Pyrimidine biosynthesis; Transferase.  
FT NON\_TER 19  
SQ SEQUENCE 19 AA; 2099 MW; 7F9B77FE83B59E21 CRC64;

Query Match 26.1%; Score 23; DB 1; Length 19;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 RQLQLSISSCLQ 16  
| | | | |  
Db 8 RPLQLNAQQLQ 19

RESULT 5  
ANCR\_AGKBI ANCR\_AGKBI STANDARD; PRT; 20 AA.  
AC P33588;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE ANCR0D (EC 3.4.21.74) (VENOMBIN A) (PROTEIN C ACTIVATOR) (ACC-C)  
DE (FRAGMENT).  
OS Agkistrodon bilineatus (Cantill) (Tropical moccasin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Viperidae; Crotalinae; Agkistrodon.  
OX NCBI\_TaxID=8718;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=90350102; PubMed=2385829;  
RA Nakagaki T., Kazim A.L., Kiesel W.;  
RT "Isolation and characterization of a protein C activator from  
RT tropical moccasin venom."  
RL Thromb. Res. 58:593-602(1990).  
CC -!- FUNCTION: THROMBIN-LIKE SNAKE VENOM SERINE PROTEASE. CLEAVES  
CC FIBRINOPEPTIDES AM, AO, AND AV; THE ABERRANT FIBRINOGEN IS THEN  
CC INCAPABLE OF BEING CROSS-LINKED, FORMING EASILY DISPERSIBLE  
CC CLOTS.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY. SNAKE VENOM SUBFAMILY.  
CC PIR; A60489; A60489.  
DR MEROPS; S01.178;  
DR InterPro; IPR001254;  
DR PROSITE; PS00134; TRYPsin\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPsin\_SER; PARTIAL.  
KW Hydrolase; Serine protease; Venom; Glycoprotein.  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2191 MW; 6E99F8B4CC53EF61 CRC64;

Query Match 25.0%; Score 22; DB 1; Length 20;  
Best Local Similarity 57.1%; Pred. No. 6e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 DHRQLQL 9  
| | | | |  
Db 11 EHRSLAL 17

RESULT 6  
RP7L\_STRSQ RP7L\_STRSQ STANDARD; PRT; 21 AA.  
AC P37046;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE RP 71955.  
OS Streptomyces sp. (strain SP9440).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=1931;  
RN [1]  
RP SEQUENCE, AND STRUCTURE BY NMR.  
RX MEDLINE=94114512; PubMed=8286361;  
RA Frechet D., Guittion J.D., Herman F., Faucher D., Helynck G.,  
RA Monegier du Sorbier B., Ridoux J.P., James-Surcouf E., Vuilhorgne M.;  
RT "Solution structure of RP 71955, a new 21 amino acid tricyclic  
RT peptide active against HIV-1 virus."  
RL Biochemistry 33:42-50(1994).  
CC -!- FUNCTION: ACTIVE AGAINST HIV-1 VIRUS.  
CC -!- PTM: THERE IS AN AMIDE BOND BETWEEN CYS-1 AND ASP-9.  
DR PDB; 1RPB; 30-APR-94.  
DR PDB; 1RPC; 30-APR-94.  
KW Antiviral; 3D-structure.  
FT DISULFID 1 13  
FT DISULFID 7 19  
FT STRAND 2 2  
FT TURN 11 12  
FT STRAND 15 15  
SQ SEQUENCE 21 AA; 2186 MW; EF94FAA21634FF91 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 21;  
Best Local Similarity 66.7%; Pred. No. 9.5e+02;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 LSISSC 14  
| | | | |  
Db 2 LGIGSC 7

RESULT 7  
FLB3\_TREHY FLB3\_TREHY STANDARD; PRT; 25 AA.  
AC P80161;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE FLAGELLAR FILAMENT CORE PROTEIN FLAB3 (32 KDA CORE PROTEIN)  
DE (FRAGMENT).  
GN FLAB3.  
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).  
OC Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.  
OX NCBI\_TaxID=159;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=C5;  
RX MEDLINE=93139764; PubMed=1487733;  
RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,  
RA van der Zeijst B.A.M., Kusters J.G.;  
RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are  
RT composed of two sheath proteins and three core proteins."  
RL J. Gen. Microbiol. 138:2697-2706(1992).  
CC -!- FUNCTION: COMPONENT OF THE CORE OF THE FLAGELLA.  
CC -!- SUBUNIT: OUTER LAYER IS COMPOSED OF TWO SHEATH PROTEINS, FLAAL  
CC (44 KDA) AND FLAAL2 (35 KDA) AND A CORE THAT CONTAINS THREE  
CC PROTEINS FLAB1 (37 KDA), FLAB2 (34 KDA) AND FLAB3 (32 KDA).  
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC FLAGELLUM.  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.  
DR PIR; D47689; D47689.  
KW Flagella; Periplasmic.  
FT NON\_TER 25  
SQ SEQUENCE 25 AA; 2746 MW; 1AAA0CA722EE3D0A CRC64;

Query Match 23.9%; Score 21; DB 1; Length 25;  
Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 HRQLQLSISS 13  
| | | | |  
Db 13 NRQLNLGTGS 22

```

RESULT 8
OXYT_RAJCL STANDARD; PRT; 9 AA.
ID AC P42994;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUMITOCIN.
OS Raja clavata (Thornback ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosqualea; Pristiogalea; Batoidea;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7781;
RN [1]
RP SEQUENCE.
RX MEDLINE=66123415; PubMed=5880565;
RA Acher R., Chauvet J., Chauvet M.-T., Crepy D.;
RT "Phylogeny of neurophyseal peptides: isolation of a new hormone,
RT glumitocin (Ser 4-Gln 8-oxytocin) present in a cartilaginous fish,
RT the ray (Raja clavata).";
RL Biochim. Biophys. Acta 107:393-396(1965).
CC -!- SIMILARITY: ANTI-DIURETIC HORMONE.
CC -!- FUNCTION: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR InterPro: IPR000981;
DR Pfam: PF00220; hormone4; 1.
DR PROSITE: PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Amidation.
FT DISULFID 1 6
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 984 MW; 17E9C76EB455B04B CRC64;

Query Match 22.7%; Score 20; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 9.3e+04;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 ISSCLQ 16
DB 3 ISNCPQ 8

RESULT 9
PROP_SALTY STANDARD; PRT; 17 AA.
ID AC P40862;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROLINE/BETAINE TRANSPORTER (PROLINE PORTER II) (PPII). (FRAGMENT).
GN PROP.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2;
RX MEDLINE=93308095; PubMed=8391535;
RA Roland K.L., Martin L.E., Esther C.R., Spitznagel J.K.;
RT "Spontaneous pmrA mutants of Salmonella typhimurium LT2 define a new
RT two-component regulatory system with a possible role in virulence.";
RL J. Bacteriol. 175:4154-4164(1993).
CC -!- FUNCTION: STRETCH-INACTIVATED PROLINE/BETAINE TRANSPORTER. PROP IS
CC BOTH AN OSMOSENSOR AND AN OSMOREGULATOR WHICH IS AVAILABLE TO
CC PARTICIPATE EARLY IN THE BACTERIAL OSMOREGULATORY RESPONSE.
CC MEDIATES THE ACTIVE ACCUMULATION OF SOLUTES SUCH AS PROLINE,
CC GLYCINE BETAINE, STACHYDRINE, PIPECOLIC ACID, ECTOINE AND TAURINE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L13395; ; NOT\_ANNOTATED\_CDS.

DR StyGene; SGI0510; prop.

DR InterPro; IPR001066; .

DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; PARTIAL.

DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; PARTIAL.

KW Transport; Transmembrane; Inner membrane; Symport.

FT NON\_TER 1 1

FT DOMAIN <1 17 CYTOPLASMIC (POTENTIAL).

SQ SEQUENCE 17 AA; 2102 MW; D3D0DB988FA3130C CRC64;

Query Match 22.7%; Score 20; DB 1; Length 17;

Best Local Similarity 45.5%; Pred. No. 1.1e+03;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 LQLSISSCLQ 17

DB 1 LQVKRSRLVQQ 11

RESULT 10

APC1\_MACFA

ID APC1\_MACFA STANDARD; PRT; 22 AA.

AC P18657;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE APOLIPOPROTEIN C-I (APO-CI) (FRAGMENT).

GN APC1.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Macaca.

OX NCBI\_TaxID=9541;

RN [1]

RP SEQUENCE.

RX MEDLINE=87185451; PubMed=3105581;

RA Herbert P.N., Bausserman L.L., Lynch K.M., Saritelli A.L.,

RA Kantor M.A., Nicolosi R.J., Shulman R.S.;

RT "Homologues of the human C and A apolipoproteins in the Macaca

RT fascicularis (cynomolgus) monkey.;"

RL Biochemistry 26:1457-1463(1987).

CC -!- FUNCTION: APPEARS TO MODULATE THE INTERACTION OF APOE WITH BETA-

CC MIGRATING VLDL AND INHIBIT BINDING OF BETA-VLDL TO THE LDL

CC RECEPTOR-RELATED PROTEIN.

CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -!- TISSUE SPECIFICITY: SECRETED IN PLASMA.

CC -!- MISCELLANEOUS: APO-CI MAKES UP ABOUT 10% OF THE PROTEIN OF THE

CC VLDL (VERY LOW DENSITY LIPOPROTEIN) & 2% OF THAT OF HDL (HIGH

CC DENSITY LIPOPROTEIN).

CC -!- SIMILARITY: BELONGS TO THE APOC1 FAMILY.

DR PIR; C26627; C26627.

DR HSP; P02654; IOPP.

KW Plasma; Lipid transport; VLDL.

FT NON\_TER 22 22

SQ SEQUENCE 22 AA; 2349 MW; 4E478A03C91EA914 CRC64;

Query Match 22.7%; Score 20; DB 1; Length 22;

Best Local Similarity 50.0%; Pred. No. 1.5e+03;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 ISSCLQQL 18

DB 4 VSSALDKL 11

```

*
RESULT 11
CH60_BOVIN STANDARD; PRT; 22 AA.
AC P31081;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 60 KDA HEAT SHOCK PROTEIN, MITOCHONDRIAL (HSP60) (50 KDA CHAPERONIN)
DE (CPN60) (HEAT SHOCK PROTEIN 60) (HSP-60) (MITOCHONDRIAL MATRIX PROTEIN
DE P1) (FRAGMENT).
DE HSPD1 OR HSP60.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE-Adrenal gland;
RX MEDLINE=93208180; PubMed=8096152;
RA Dai Z., Lackland H., Stein S., Li Q., Radziewicz R., Williams R.,
RA Sigal L.H.;
RT "Molecular mimicry in Lyme disease: monoclonal antibody H9724 to B.
RT burgdorferi flagellin specifically detects chaperonin-HSP60.";
RL Biochim. Biophys. Acta 1181:97-100(1993).
CC -!- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
CC MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
CC IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
CC REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
CC UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR HSP; P06139; IAOB.
DR InterPro; IPR001844; .
DR PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
KW Chaperone; ATP-binding; Mitochondrion.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2303 MW; E09D2EB9934475F5 CRC64;

```

```

Query Match 22.7%; Score 20; DB 1; Length 22;
Best Local Similarity 62.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 ADHRQLQL 9
   | | | |
DB 8 ADARLML 15

```

```

RESULT 12
ODPB_BOVIN STANDARD; PRT; 24 AA.
AC P11966;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PYRUVATE DEHYDROGENASE E1 COMPONENT BETA SUBUNIT (EC 1.2.4.1)
DE (FRAGMENT).
GN PDHB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=88134251; PubMed=2829898;
RA Ho L., Javed A.A., Pepin R.A., Thekkumkara T.J., Raefsky C.,
RA Mole J.E., Caliendo A.M., Kwon M.S., Kerr D.S., Patel M.S.;
RT "Identification of a cDNA clone for the beta-subunit of the pyruvate
RT dehydrogenase component of human pyruvate dehydrogenase complex.";
RL Biochem. Biophys. Res. Commun. 150:904-908(1988).
CC -!- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL

```

```

CC CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE
CC COPIES OF THREE ENZYMIC COMPONENTS: PYRUVATE DEHYDROGENASE (E1),
CC DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2) & LIPOAMIDE DEHYDROGENASE
CC (E3).
CC -!- CATALYTIC ACTIVITY: PYRUVATE + LIPOAMIDE -> S-Acetyl-Dihydro-
CC LIPOAMIDE + CO(2).
CC -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC -!- SUBUNIT: Tetramer of two alpha and two beta subunits.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
DR PIR; B27712; B27712.
KW Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
KW Mitochondrion.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2849 MW; 0E9DF43248221854 CRC64;

```

```

Query Match 22.7%; Score 20; DB 1; Length 24;
Best Local Similarity 27.3%; Pred. No. 1.7e+03;
Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 7 LQLSISSCLOQ 17
   | | | | |
DB 1 LQTVREAINQ 11

```

```

RESULT 13
RECJ_SALTY STANDARD; PRT; 14 AA.
AC P28355;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE SINGLE-STRANDED-DNA-SPECIFIC EXONUCLEASE RECJ (EC 3.1.-.-) (FRAGMENT).
GN RECJ.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91046011; PubMed=2236050;
RA Kawakami K., Nakamura Y.;
RT "Autogenous suppression of an opal mutation in the gene encoding
RT peptide chain release factor 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8432-8436(1990).
CC -!- FUNCTION: SINGLE-STRANDED-DNA-SPECIFIC EXONUCLEASE. REQUIRED FOR
CC MANY TYPES OF RECOMBINATIONAL EVENTS, ALTHOUGH THE STRINGENCY OF
CC THE REQUIREMENT FOR RECJ APPEARS TO VARY WITH THE TYPE OF
CC RECOMBINATIONAL EVENT MONITORED AND THE OTHER RECOMBINATION GENE
CC PRODUCTS WHICH ARE AVAILABLE.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M38590; AAA72913.1; .
CC StyGene; SG10332; recj.
KW Hydrolase; Nuclease; Exonuclease.
FT NON_TER 1
SQ SEQUENCE 14 AA; 1696 MW; 105E784AC26C5650 CRC64;

```

```

Query Match 21.6%; Score 19; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4 HRQLQLSI 11
   | | | |
DB 1 NRSLQIII 8

```

RL Mol. Genet. 218:536-544(1989).  
CC -!- FUNCTION: MAY PLAY A ROLE IN A REDOX PROCESS INVOLVED IN NITROGEN  
CC FIXATION.  
CC -!- SUBUNIT: FIXA AND FIXB FORM A HETERODIMER (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE ETF BETA-SUBUNIT / FIXA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X16521; CAA34527.1; -  
DR PIR; PQ0022; PQ0022.  
DR InterPro; IPR000049; -  
DR PROSITE; PS01065; ETF\_BETA; PARTIAL.  
KW Electron transport; Nitrogen fixation.  
FT NON\_TER 18 18  
SQ SEQUENCE 18 AA; 2036 MW; 74973C8BA2087663 CRC64;  
  
Query Match 21.6%; Score 19; DB 1; Length 18;  
Best Local Similarity 30.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
  
QY 9 LSISSCLOQL 18  
: | | | |  
Db 1 MHIVVCIQV 10  
  
Search completed: May 8, 2001, 15:14:06  
Job time: 527 sec

RESULT 14  
TPIS\_PIMPS STANDARD; PRT; 17 AA.  
AC P81666;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE TRIOSPHOSPHATE ISOMERASE, CYTOSOLIC (EC 5.3.1.1) (TIM) (FRAGMENTS).  
OS Pinus pinaster (Maritime pine).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_TaxID=71647;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Needle;  
RX MEDLINE=99274086; PubMed=10344291;  
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,  
RA Frigerio J.-M., Plomion C.;  
RT "Separation and characterization of needle and xylem maritime pine  
proteins";  
RL Electrophoresis 20:1098-1108(1999).  
CC -!- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE = DIHYDROXY-  
CC ACETONE PHOSPHATE.  
CC -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
CC -!- INDUCTION: BY WATER-STRESS.  
CC -!- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC  
CC AND PLASTID.  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN  
CC (SPOT N139) IS: 5.9, ITS MW IS: 24 KDA.  
CC -!- SIMILARITY: BELONGS TO THE TRIOSPHOSPHATE ISOMERASE FAMILY.  
DR InterPro; IPR000652; -  
DR PROSITE; PS00171; TIM; PARTIAL.  
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;  
KW Pentose shunt; Multigene family.  
FT NON\_TER 1 1  
FT NON\_CONS 9 10  
FT NON\_TER 17 17  
SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;  
  
Query Match 21.6%; Score 19; DB 1; Length 17;  
Best Local Similarity 60.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 14 CLQQL 18  
: | | | |  
Db 5 CYEQL 9  
  
RESULT 15  
FIXA\_RHILE STANDARD; PRT; 18 AA.  
AC P14313;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE FIXA PROTEIN (FRAGMENT).  
GN FIXA.  
OS Rhizobium leguminosarum.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=384;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90066358; PubMed=2555670;  
RA Hontelez J.G.J., Lankhorst R.K., Katinakis P., van den Bos R.C.,  
RA van Rammem A.;  
RT "Characterization and nucleotide sequence of a novel gene fixW  
upstream of the fixABC operon in Rhizobium leguminosarum.";



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:16:04 ; Search time 114.89 Seconds  
(without alignments)  
18.363 Million cell updates/sec

Title: US-09-165-546A-8

Perfect score: 88

Sequence: 1 AADHRQLQLSISSCLQL 18

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_15.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organalle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_unclassified.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	28	31.8	18	13	O57600 gallus gall
2	26	29.5	24	2	O54460 erwinia her
3	26	29.5	24	5	O9u922 lingua ana
4	26	29.5	25	5	O25185 artemia sal
5	26	29.5	25	13	O9PWR9 xiphophorus
6	25	28.4	17	10	O9S919 petunia hyb
7	24	27.3	12	11	O9Q2Y4 mus musculu
8	24	27.3	17	5	O9TVX7 bombyx mori
9	24	27.3	20	14	O64976 alfalfa mos
10	24	27.3	21	5	O25084 hermania m
11	24	27.3	22	13	O91103 morone saxa
12	24	27.3	22	13	O91106 morone saxa
13	24	27.3	22	13	O91107 morone saxa
14	24	27.3	23	5	O25139 haliotis ru
15	24	27.3	25	5	O23965 dendrocoelu
16	24	27.3	25	5	O25872 polycelis f
17	23.5	26.7	25	10	O9S8V9 triticum ae
18	23	26.1	13	2	O53313 corynebacte
19	23	26.1	14	11	O9Z1H4 mus musculu

20	23	26.1	17	2	O45304
21	23	26.1	22	13	O91104
22	23	26.1	22	13	O91105
23	23	26.1	22	13	O91108
24	23	26.1	24	8	O79910
25	23	26.1	25	11	O63986
26	23	26.1	25	11	O63987
27	23	26.1	25	11	O63991
28	22	25.0	15	3	O9UR64
29	22	25.0	17	5	O17203
30	22	25.0	19	2	O51076
31	22	25.0	19	13	P81882
32	22	25.0	20	8	O34694
33	22	25.0	21	2	O25126
34	22	25.0	21	2	O25621
35	22	25.0	21	2	O9ZKB8
36	22	25.0	22	6	O9TR02
37	22	25.0	23	5	O25133
38	22	25.0	24	4	O15133
39	22	25.0	24	6	O9TRL9
40	22	25.0	25	13	O9YGG0
41	22	25.0	25	13	O9PWS0
42	21.5	24.4	16	4	O00497
43	21.5	24.4	22	6	O9TU54
44	21	23.9	17	2	O34216
45	21	23.9	18	13	O9PRM8

ALIGNMENTS

RESULT 1			
O57600	PRELIMINARY;	PRT;	18 AA.
ID	O57600		
AC	O57600;		
DT	01-JUN-1998 (TReMBLrel. 06, Created)		
DT	01-JUN-1998 (TReMBLrel. 06, Last sequence update)		
DT	01-AUG-1998 (TReMBLrel. 07, Last annotation update)		
DE	THROMBOMUCIN (FRAGMENT).		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	McNagny K.M., Petterson I., Rossi F., Flamme I., Shevchenko A.,		
RA	Mann M., Graf T.;		
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; Y13976; CAA74309.1; -.		
FT	NON_TER 1 1		
SQ	SEQUENCE 18 AA; 1973 MW; CA71002CD77FE4F9 CRC64;		

Query Match	31.8%;	Score 28;	DB 13;	Length 18;
Best Local Similarity	55.6%;	Pred. No. 2.4e+02;		
Matches	5;	Conservative	2;	Mismatches
			2;	Indels
				0;
				Gaps
				0;
QY	8	QLSISSCLQ 16		
		I: : l l		
Db	7	QVDAASCLQ 15		

RESULT 2			
O54460	PRELIMINARY;	PRT;	24 AA.
ID	O54460		
AC	O54460;		
DT	01-JUN-1998 (TrEMBLrel. 06, Created)		
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)		
DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)		
DE	HYPOTHETICAL 2.7 KDA PROTEIN (FRAGMENT).		
OS	Erwinia herbicola.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		

```

OC Erwinia.
OX NCBI_TaxID=557;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98083064; PubMed=9422601;
RT Subramanian P.S., Xie G., Xia T., Jensen R.A.;
RT "Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate
RT synthase from Neisseria gonorrhoeae in the context of its membership
RT in a protein family containing a subset of 3-deoxy-D-arabino-
RT heptulosonate 7-phosphate synthases.";
RL J. Bacteriol. 180:119-127(1998).
DR EMBL; U93355; AAB96401.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2725 MW; 34989DD77385CEE CRC64;

Query Match 29.5%; Score 26; DB 2; Length 24;
Best Local Similarity 38.1%; Pred. No. 7.2e+02;
Matches 8; Conservative 4; Mismatches 1; Indels 8; Gaps 1;

QY 1 AADHRQ-----LQLSISS 13
|:::| |::|
Db 4 AAEYRQPPDPWGKTLVSCSS 24

RESULT 3
QY0922 PRELIMINARY; PRT; 24 AA.
AC QY0922;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE LOX2 HOMEODOMAIN PROTEIN (FRAGMENT).
OS Lingula anatina.
OC Eukaryota; Metazoa; Brachiopoda; Linguliformes; Lingulata; Lingulida;
OC Linguloidea; Lingulidae; Lingula.
OX NCBI_TaxID=7574;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99318125; PubMed=10391241;
RA de Rosa R., Grenier J.K., Andreeva T., Cook C.E., Adoutte A., Akam M.,
RA Carroll S.B., Balavoine G.;
RT "Hox genes in brachiopods and priapulids and protostome evolution.";
RL Nature 399:772-776(1999).
DR EMBL; AF144677; AAD45592.1; -
DR HSSP; P02833; IHOM.
DR INTERPRO; IPR001356; -
DR PFAM; PF00046; homeobox; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1
FT NON_TER 24
SQ SEQUENCE 24 AA; 3106 MW; 1C57BBI3E6AF1524 CRC64;

Query Match 29.5%; Score 26; DB 5; Length 24;
Best Local Similarity 38.5%; Pred. No. 7.2e+02;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 RQLQLSISSCLOQ 17
|:::| |::|
Db 9 RRIELSHMLCLTE 21

RESULT 4
Q25185 PRELIMINARY; PRT; 25 AA.
AC Q25185;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ANTENNAPEDIA CLASS HOMEBOX-RELATED PROTEIN (FRAGMENT).
GN CNH4.

```

```

OS Artemia salina (Brine shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
OX NCBI_TaxID=85549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93321609; PubMed=8101168;
RT Shenk M.A., Bode H.R., Steele R.E.;
RT "Expression of Cox-2, a HOM/HOX homeobox gene in hydra, is correlated
RT with axial pattern formation.";
RL Development 117:657-667(1993).
DR EMBL; M62873; AAC83403.1; -
DR HSSP; P02833; IHOM.
DR INTERPRO; IPR001356; -
DR PFAM; PF00046; homeobox; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3127 MW; 1AB483ADC8A3BFCF CRC64;

Query Match 29.5%; Score 26; DB 5; Length 25;
Best Local Similarity 30.8%; Pred. No. 7.5e+02;
Matches 4; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 RQLQLSISSCLOQ 17
|:::| |::|
Db 9 RRIEMAHSLCLTE 21

RESULT 5
QY09WR9 PRELIMINARY; PRT; 25 AA.
AC QY09WR9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE MELANOMA RECEPTOR TYROSINE KINASE (FRAGMENT).
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8083;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ONC SR-STRAIN;
RA Scharf M., Wilde B., Hornung U.;
RT "Triplet repeat variability in the signal peptide sequence of the Xmrk
RT receptor tyrosine kinase gene in Xiphophorus fish.";
RL Gene 224:17-21(1998).
DR EMBL; U82804; AAD10123.1; -
KW Receptor; Kinase.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2695 MW; 967AF74362DF4350 CRC64;

Query Match 29.5%; Score 26; DB 13; Length 25;
Best Local Similarity 75.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LQLSISSC 14
|:::| |::|
Db 15 LVLSISRC 22

RESULT 6
QY09S19 PRELIMINARY; PRT; 17 AA.
AC QY09S19;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)

```

DE SO-PROTEIN-31 KDA.  
OS Petunia hybrida (Petunia).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
OC Solanales; Solanaceae; Petunia.  
OX NCBI\_TaxID=4102;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92114864; PubMed=1766433;  
RA Ai Y.J., Kron E., Kao T.H.;  
RT "S-alles are retained and expressed in a self-compatible cultivar of  
Petunia hybrida";  
RL Mol. Gen. Genet. 230:353-358(1991).  
SQ SEQUENCE 17 AA; 2035 MW; 59270B606B574F63 CRC64;

Query Match 28.4%; Score 25; DB 10; Length 17;  
Best Local Similarity 60.0%; Pred. No. 8.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AADHRQLQLS 10  
| | | | |  
Db 1 AFDHWQLVLT 10

RESULT 7  
Q9Q2Y4  
ID Q9Q2Y4 PRELIMINARY; PRT; 12 AA.  
AC Q9Q2Y4;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE PROTEIN KINASE LKBL (FRAGMENT).  
GN LKBL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129;  
RX MEDLINE=99330555; PubMed=10400995;  
RA Smith D.P., Spicer J., Smith A., Swift S., Ashworth A.;  
RT "The mouse peutz-jeghers syndrome gene lkb1 encodes a nuclear protein  
kinase";  
RL Hum. Mol. Genet. 8:1479-1485(1999).  
DR EMBL; AF145697; AAD55369.1;  
KW Kinase.  
FT NON\_TER 1  
SQ SEQUENCE 12 AA; 1445 MW; 6B06A6C54BC5B734 CRC64;

Query Match 27.3%; Score 24; DB 11; Length 12;  
Best Local Similarity 57.1%; Pred. No. 9e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 ISSCLQ 17  
: | : | |  
Db 6 LSACKQ 12

RESULT 8  
Q9TVX7  
ID Q9TVX7 PRELIMINARY; PRT; 17 AA.  
AC Q9TVX7;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE A.X3 PROTEIN (FRAGMENT).  
GN A.X3 OR A.L4 OR A.L13 OR A.R2 OR A.R3 OR A.X1 OR A.X2.  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;

OC Bombycoidea; Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=703;  
RX MEDLINE=90040707; PubMed=2810362;  
RA Spoerel N.A., Nguyen H.T., Eickbush T.H., Kafatos F.C.;  
RT "Gene evolution and regulation in the chorion complex of Bombyx mori.  
Hybridization and sequence analysis of multiple developmentally middle  
A/B chorion gene pairs";  
RL J. Mol. Biol. 209:1-19(1989).  
DR EMBL; X15572; CAA33595.1;  
DR EMBL; X15560; CAA33571.1;  
DR EMBL; X15566; CAA33583.1;  
DR EMBL; X15568; CAA33587.1;  
DR EMBL; X15569; CAA33589.1;  
DR EMBL; X15570; CAA33591.1;  
DR EMBL; X15571; CAA33593.1;  
FT NON\_TER 17  
SQ SEQUENCE 17 AA; 1901 MW; 5FD3450433955C8C CRC64;

Query Match 27.3%; Score 24; DB 5; Length 17;  
Best Local Similarity 45.5%; Pred. No. 1.2e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 LQSISSCLQ 17  
| | : | | |  
Db 7 LLCVQACLIQ 17

RESULT 9  
Q64976  
ID Q64976 PRELIMINARY; PRT; 20 AA.  
AC Q64976;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE 115 KDA PROTEIN (FRAGMENT).  
OS Alfalfa mosaic virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;  
OC Alfamovirus.  
OX NCBI\_TaxID=12321;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ALMV-S;  
RX MEDLINE=83220821; PubMed=6856476;  
RA Ravelonandro M., Godefroy-Colburn T., Pinck L.;  
RT "Structure of the 5'-terminal untranslated region of the genomic RNAs  
from two strains of alfalfa mosaic virus";  
RL Nucleic Acids Res. 11:2815-2826(1983).  
DR EMBL; M35975; AAA66593.1;  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2178 MW; E66DC56E6E7687A9 CRC64;

Query Match 27.3%; Score 24; DB 14; Length 20;  
Best Local Similarity 71.4%; Pred. No. 1.4e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 LSISLCL 15  
| | | | |  
Db 6 LSTDSC 12

RESULT 10  
Q25084  
ID Q25084 PRELIMINARY; PRT; 21 AA.  
AC Q25084;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE CLONE AH0X2 HOMEBOX PROTEIN (FRAGMENT).

OS	Herdmania momus.
OC	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OX	Stolidobranchia; Pyuridae; Herdmania.
OX	NCBI_TaxID=7733;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=EMBRYO;
RA	Kennett C.V.D.;
RL	Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.
RR	EMBL; U09802; AAA18627.1; -.
DR	HSSP; P02833; IHOM.
DR	INTERPRO: IPR001356; -.
DR	PFAM: PF00046; homeobox; 1.
KW	Homeobox; DNA-binding; Nuclear protein.
FT	NON_TER 1
FT	NON_TER 21
SQ	SEQUENCE 21 AA: 2627 MW: 5DCDB3BF1556500C CRC64:

```
Query Match      27.3% Score 24; DB 5; Length 21;
Best Local Similarity . 23.1%; Pred. NO. 1.5e+03;
Matches 3; Conservative 7; Mismatches 3; Indels
```

---

```
QY    5 RQLQLSISSCLOQ 17  
      |:::~::~|||  
Db     8 RRRTIAHALCLSE 20
```

RESULT	11
Q91103	
ID	PRELIMINARY; PRT; 22 AA.
AC	Q91103;
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	HOK-A4-LIKE HOMEDOMAIN PROTEIN (FRAGMENT).
OS	Morone saxatilis (Striped bass).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC	Moronidae; Morone.
OX	NCBI_TaxID=34816;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BLOOD;
RX	MEDLINE=95005122; PubMed=7921046;
RA	Pavell A.M., Stellingwag E.J.;
RT	"Survey of Hox-like genes in the teleost Morone saxatilis:
RT	implications for evolution of the Hox gene family.";
RL	Mol. Mar. Biol. Biotechnol. 3:149-157(1994).
DR	EMBL; U09945; AAC59651.1; -
DR	HSSP; P02833; LHOM.
DR	INTERPRO; IPR001356; -
DR	PFAM; PF00046; homeobox: 1.
KW	Homeobox; Nuclear protein; DNA-binding.
FT	NON_TER 1
FT	NON_TER 22
SQ	SEQUENCE 22 AA; 2733 MW; 37594933CDB3D7B4 CRC64;

```

Query Match      27.3%; Score 24; DB 13; Length 22;
Best Local Similarity 23.1%; Pred. No. 1.5e+03;
Matches 3; Conservative 7; Mismatches 3; Indels
QY 5 RQLQLSISSLQ 17
|::: : || :
Db 6 RRIEANTLCLE 18

```

RESULT 12	
Q91106	
ID Q91106	PRELIMINARY;
AC Q91106;	PRT; 22 AA.

DT	01-NOV-1996	(TrEMBLrel. 01, Created)	
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)	
DT	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)	
DE	HGX-B5-LIKE HOMEODOMAIN PROTEIN (FRAGMENT).		
OS	Morone saxatilis (Striped bass).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;		
OC	Moronidae; Morone.		
OX	NCBI_TaxID=34816;		
	[1]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=BLOOD;		
RC			
RX	MEDLINE=95005122; Pubmed=7921046;		
RA	Pavell A.M., Stellwag E.J.;		
RT	"Survey of Hox-like genes in the teleost Morone saxatilis:		
RT	Implications for evolution of the Hox gene family.";		
RL	Mol. Mar. Biol. Biotechnol. 3:149-157(1994).		
DR	EMBL; U09948; AAC59654.1; -.		
DR	HSSP; P02833; IHOM.		
DR	INTERPRO; IPR001356; -.		
DR	PFAM; PF00046; homeobox; 1.		
KW	Homeobox; Nuclear protein; DNA-binding.		
FT	NON_TER	1	
FT	NON_TER	22	
FT	SEQUENCE	22 AA; 2726 MW;	37494DADCDB3D7B4 CRC64;
SQ			

Query Match 27.3%; Score 24; DB 13; Length 22;  
Best Local Similarity 23.1%; Pred. No. 1.5e+03;  
Matches 3; Conservative 7; Mismatches 3; Indels

QY	5	ROLQLSISSCLQQ	17
		::: :	
Db	6	RRIEIAHALCLSE	18

RESULT 13  
091107

ID	Q91107	PRELIMINARY;	PRT;	22 AA.
AC	Q91107;			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)		
DT	01-MAY-2000	(TREMBLrel. 13, Last annotation update)		
DE	HOX-C4-LIKE HOMEODOMAIN PROTEIN (FRAGMENT).			
OS	Morone saxatilis (Striped bass).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthopterygii; Acanthopterygii; Percomorpha; Perciformes; Percoidei;			
OC	Moronidae; Morone.			
OX	NCBI_TaxID=34816;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BLOOD;			
RC	MEDLINE=95005122; Pubmed=7921046;			
RA	Pavell A.M., Stellwag E.J.;			
RT	"Survey of Hox-like genes in the teleost Morone saxatilis:			
RT	Implications for evolution of the Hox gene family.;"			
RL	Mol. Mar. Biol. Biotechnol. 3:149-157(1994).			
DR	ENBL; U09949; AAC59655.1; "			
DR	INTERPRO; IPR001356; "			
DR	PRAM; PF000046; homeobox; 1			
KW	Homeobox; Nuclear protein; DNA-binding.			
FT	NON_TER	1		
FT	NON_TER	22		
SQ	SEQUENCE	22 AA;	2726 MW;	37494DADCDB3D7B4 CRC64;

Query Match 27.3%; Score 24; DB 13; Length 22;  
Best Local Similarity 23.1%; Pred. No. 1.5e+03;  
Matches 3; Conservative 7; Mismatches 3; Indels

QY . 5 RQLQLSISSCLQQ 17





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:07:07 ; Search time 113.31 Seconds  
(without alignments)  
9.081 Million cell updates/sec

Title: US-09-165-546A-9  
Perfect score: 84  
Sequence: 1 VLLKEFTVSGNLTIRLT 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 165078

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_0401.\*  
1: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	81.5	17	Y52436	Human tumour antiq
2	56	66.7	18	Y52440	Human tumour antiq
3	48	57.1	10	Y06001	Human cancer antiq
4	48	57.1	19	Y52437	Human tumour antiq
5	47	56.0	10	Y05988	Human cancer antiq
6	47	56.0	10	Y05988	Human cancer antiq
7	44	52.4	9	Y79753	NY-ESO-1 derived p
8	42	50.0	9	Y06053	Human cancer antiq
9	42	50.0	9	Y06030	Human cancer antiq
10	42	50.0	9	Y06036	Human cancer antiq
11	42	50.0	9	Y79758	NY-ESO-1 derived p

12	37	44.0	9	20	Y06027	Human cancer antiq
13	37	44.0	9	20	Y06028	Human cancer antiq
14	37	44.0	10	20	Y06000	Human cancer antiq
15	35	41.7	25	19	W39953	Peptide effecting
16	33	39.3	15	21	Y93064	Transforming grow
17	31	36.9	9	20	Y06034	Human cancer antiq
18	31	36.9	20	11	R04057	Reagent of GTP-bin
19	29	34.5	15	21	Y93063	Transforming grow
20	29	34.5	20	21	B13701	C. pneumoniae Cp-S
21	28	33.3	12	14	R35239	D32.39 antibody is
22	28	33.3	12	15	R56707	Random peptide #57
23	28	33.3	12	17	R91455	D32.39 monoclonal
24	28	33.3	12	18	W25237	Antibody D32.39 ep
25	28	33.3	20	19	W42135	T-cell epitope pep
26	28	33.3	21	13	R27159	LFA-3 CD2 binding
27	28	33.3	23	21	Y64835	Human 5' EST relat
28	28	33.3	24	18	W33935	Betal-adrenergic r
29	28	33.3	24	20	Y41575	Fragment of human
30	27	32.1	10	17	R96042	Telomerase peptide
31	27	32.1	10	20	Y06017	Human cancer antiq
32	27	32.1	20	20	Y43182	S. rochei strain E
33	27	32.1	21	15	R62128	U1 snRNP 70K prote
34	27	32.1	22	18	W08416	Exosite III from h
35	27	32.1	24	21	Y56838	Apple LRPkml LRR s
36	27	32.1	25	18	W33944	Alpha1A-adrenergic
37	26	31.0	14	18	W36768	Thrombopoietin rec
38	26	31.0	14	18	W36687	Thrombopoietin rec
39	26	31.0	14	18	W09536	Thrombopoietin rec
40	26	31.0	15	17	R97311	Parietaria-based s
41	26	31.0	17	20	Y02822	Fragment of human
42	26	31.0	18	18	W27513	Human interleukin
43	26	31.0	21	21	Y98333	Alpha D peptide de
44	26	31.0	22	21	Y98384	Alpha D peptide de
45	26	31.0	25	17	R97310	Parietaria-based s

## ALIGNMENTS

RESULT 1  
Y52436  
ID Y52436 standard; Protein; 17 AA.

XX Y52436;

XX AC

XX 15-FEB-2000 (first entry)

XX XX Human tumour antigen.NY-ESO-1 peptide #9.

XX Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;

XX T-cell; helper; stimulation; proliferation; treatment;

XX diagnosis; prevention; melanoma; breast cancer; ovarian cancer;

XX prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;

XX lymphoma.

XX Synthetic.

XX Homo sapiens.

XX OS

XX W09953938-A1.

XX PN

XX 28-OCT-1999.

XX PD

XX XX

XX 24-MAR-1999; 99WO-US06875.

XX XX

XX 17-APR-1998; 98US-0062422.

XX PR

XX 02-OCT-1998; 98US-0165546.

XX XX

XX (LUDW-) LUDWIG INST CANCER RES.

XX PA

XX NY-ESO-1 derived p

XX PI

XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;

XX Gure A, Ritter G;

XX XX

XX WPI; 2000-038483/03.

XX DR

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
PT useful for therapeutic and diagnostic purposes

XX Claim 4; Page 22; 49pp; English.

XX Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
CC localisation studies revealed it to be expressed at high levels  
CC in normal ovary and testis but not in normal colon, kidney, liver,  
CC brain, oesophagus and skin. It was expressed in certain tumours and  
CC tumour cell lines with some degree of frequency - these included  
CC melanoma specimens and cell lines, and breast and bladder cancer  
CC specimens, with expression in other tumour types being sporadic.  
CC These NY-ESO-1-derived peptides may be used in methods and  
CC compositions used for the treatment, diagnosis and prevention of  
CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
CC or lymphoma) and to stimulate the proliferation of T cells.

XX Sequence 17 AA:

Query Match 81.5%; Score 68.5; DB 21; Length 17;  
Best Local Similarity 94.4%; Pred. No. 2.6e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 VLLKEFTVSGNLTIRLT 18  
||||| |||||||||  
Db 1 vllkef-vsgnltirlt 17

RESULT 2  
Y52440  
ID Y52440 standard; Protein; 18 AA.

XX AC Y52440;  
XX  
XX 15-FEB-2000 (first entry)  
XX Human tumour antigen NY-ESO-1 peptide #13.

DE Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
KW T-cell; helper; stimulation; proliferation; treatment;  
KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
KW lymphoma.

XX Synthetic.  
OS Homo sapiens.  
OS  
XX W09953938-A1.  
XX 28-OCT-1999.  
XX  
XX 24-MAR-1999; 99WO-US06875.  
XX  
XX 17-APR-1998; 98US-0062422.  
XX 02-OCT-1998; 98US-0165546.  
XX  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX  
XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
PI Gure A, Ritter G;  
XX WPI; 2000-038483/03.

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
PT useful for therapeutic and diagnostic purposes

PS Claim 4; Page 22; 49pp; English.

XX Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
CC localisation studies revealed it to be expressed at high levels  
CC in normal ovary and testis but not in normal colon, kidney, liver,  
CC brain, oesophagus and skin. It was expressed in certain tumours and  
CC tumour cell lines with some degree of frequency - these included  
CC melanoma specimens and cell lines, and breast and bladder cancer  
CC specimens, with expression in other tumour types being sporadic.  
CC These NY-ESO-1-derived peptides may be used in methods and  
CC compositions used for the treatment, diagnosis and prevention of  
CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
CC or lymphoma) and to stimulate the proliferation of T cells.

XX Sequence 18 AA:

Query Match 66.7%; Score 56; DB 21; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.0006;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 TVSGNLTIRLT 18  
||||| |||||||  
Db 1 tvsgnltirlt 12

RESULT 3  
Y06001  
ID Y06001 standard; Peptide; 10 AA.

XX AC Y06001;  
XX  
XX 16-AUG-1999 (first entry)  
XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

DE NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.  
OS  
OS W09918206-A2.  
XX  
XX 15-APR-1999.  
XX  
XX 21-SEP-1998; 98WO-US19609.  
XX  
XX 08-OCT-1997; 97US-0061428.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Rosenberg SA, Wang RF;  
PI WPI; 1999-277270/23.  
XX  
XX Cancer antigen NY ESO1/CAG-3  
PT  
XX Example 10; Page 42; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
CC screen for epitopes from the coding region of human NY ESO-1/CAG-3  
CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
CC The present peptide (ranked 14) corresponds to amino acid residues

CC 126-135 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 10 AA;

Query Match 57.1%; Score 48; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FTVSGNLTIF 15  
 | | | | | | | | | |  
 Db 1 ftvsgnlti 10

RESULT 4  
 Y52437  
 ID Y52437 standard; Protein; 19 AA.  
 AC Y52437;  
 XX  
 DT 15-FEB-2000 (first entry)  
 DE Human tumour antigen NY-ESO-1 peptide #10.  
 KW Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
 KW T-cell; helper; stimulation; proliferation; treatment;  
 KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
 KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
 KW lymphoma.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9953938-A1.  
 PD 28-OCT-1999.  
 XX  
 PF 24-MAR-1999; 99WO-US06875.  
 XX  
 PR 17-APR-1998; 98US-0062422.  
 PR 02-OCT-1998; 98US-0165546.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
 PI Gure A, Ritter G;  
 XX  
 DR WPI: 2000-038483/03.  
 XX  
 PT Novel peptides which bind to MHC class I and MHC class II molecules,  
 PT useful for therapeutic and diagnostic purposes -  
 XX  
 PS Claim 4; Page 22; 49pp; English.  
 XX  
 CC Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels  
 CC in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and

CC tumour cell lines with some degree of frequency - these included  
 CC melanoma specimens and cell lines, and breast and bladder cancer  
 CC specimens, with expression in other tumour types being sporadic.  
 CC These NY-ESO-1-derived peptides may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of  
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
 CC or lymphoma) and to stimulate the proliferation of T cells.

XX Sequence 19 AA;

Query Match 57.1%; Score 48; DB 21; Length 19;  
 Best Local Similarity 83.3%; Pred. No. 0.02;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLLKEFTVSGNI 12  
 : | | | | | | | | | |  
 Db 8 llkleftvsgni 19

RESULT 5  
 Y05988  
 ID Y05988 standard; Peptide; 10 AA.  
 AC Y05988;  
 XX  
 DT 16-AUG-1999 (first entry)  
 DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.  
 XX  
 KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.  
 XX Homo sapiens.  
 XX  
 PN WO9918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Rosenberg SA, Wang RF;  
 XX  
 DR WPI: 1999-277270/23.  
 XX  
 PT Cancer antigen NY ESO1/CAG-3  
 XX  
 PS Example 10; Page 42; 88pp; English.  
 XX  
 CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human NY ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 1) corresponds to amino acid residues  
 CC 127-136 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, kidney cancer and

CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 SQ Sequence 10 AA;

Query Match 56.0%; Score 47; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 10; Conservative. 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TVSGNLTIR 16  
 |||||  
 Db 1 tvsgnltir 10

RESULT 6  
 Y05980  
 ID Y05980 standard; Peptide; 10 AA.  
 XX  
 AC Y05980;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 ORF1 cancer peptide ESO10-127.  
 XX  
 KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Rosenberg SA, Wang RF;  
 XX  
 DR WPI; 1999-277270/23.  
 XX  
 PT Cancer antigen NY ESO1/CAG-3  
 XX  
 PS Claim 17; Page 64; 88pp; English.  
 XX  
 CC This sequence represents cancer peptide ESO10-127 that corresponds  
 CC to amino acid residues 127-136 of human ESO-1/CAG-3 (or CAG-3) ORF1  
 CC (see Y05965), a new and potent tumour antigen capable of eliciting  
 CC an antigen specific immune response by T cells. Cancer peptides  
 CC derived from CAG-3 ORF1, CAG-3 ORF2 (see Y05966), portions of them  
 CC and their variants (see Y05967-87), are useful as cancer vaccines  
 CC that protect against cancer. The invention provides: vectors and  
 CC host cells (also useful as vaccines); a method of diagnosis of  
 CC cancer or precancer; a transgenic animal; antisense oligonucleotides  
 CC that inhibit expression of the cancer peptide or tumour antigen;  
 CC antibodies reacting with a CAG-3 cancer peptide, useful in  
 CC diagnostic and detection assays; and methods for preventing or  
 CC inhibiting cancer by administering a cancer peptide, with or without  
 CC an HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers. Melanoma is treated by inducing cancer-specific T  
 CC cells in vitro for subsequent return to a patient.

XX Sequence 10 AA;

Query Match 56.0%; Score 47; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TVSGNLTIR 16  
 |||||  
 Db 1 tvsgnltir 10

RESULT 7  
 Y79753  
 ID Y79753 standard; Peptide; 9 AA.  
 XX  
 AC Y79753;  
 XX  
 DT 10-MAY-2000 (first entry)  
 XX  
 DE NY-ESO-1 derived peptide #9.  
 XX  
 KW Cancer: SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;  
 KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;  
 KW cytostatic; melanoma; synovial sarcoma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200000824-A1.  
 XX  
 PD 06-JAN-2000.  
 XX  
 PF 25-JUN-1999; 99WO-US14493.  
 XX  
 PR 26-JUN-1998; 98US-0105839.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 XX  
 PI Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;  
 PI Chen Y, Gure A, Old LJ;  
 XX  
 DR WPI; 2000-170933/15.  
 XX  
 PT Determining the possible presence of breast, endometrial, colorectal,  
 PT lung, bladder or head-neck cancer  
 XX  
 PS Example 13; Page 26; 40pp; English.  
 XX  
 CC A method has been developed for determining the possible presence of a  
 CC cancer, which is not melanoma or synovial sarcoma. The method comprises  
 CC assaying a sample taken from the subject to determine the expression of  
 CC an SSX gene, and determining the expression as a determination of the  
 CC possible presence of cancer. Expression of SSX1 gene indicates possible  
 CC presence of breast, endometrial, colorectal, lung, bladder or head-neck  
 CC cancer. These cancers are also detected by SSX2 and SSX4 gene expression.  
 CC SSX2 gene expression additionally indicates possible presence of  
 CC lymphoma, renal cell cancer, glioma and prostate cancer. Expression of  
 CC SSX4 gene also indicates possible presence of ovarian or stomach cancer.  
 CC SSX5 gene expression indicates the same cancers as SSX1, except breast  
 CC cancer. Determining expression of SSX gene can be used to monitor  
 CC progress of melanoma or synovial sarcoma, which is not cancer. The  
 CC SSX-derived peptide complex stimulates proliferation of cytolytic T  
 CC cells. This is useful for treating cancer, especially melanoma. Y78464  
 CC to Y78468 represent specifically claimed HLA binding peptides for use in  
 CC the method of the invention. Z88452 to Z88465 represent PCR primers used  
 CC in the isolation of SSX genes in the exemplification of the present  
 CC invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides  
 CC derived from SSX proteins or NY-ESO-1, which are used in the  
 CC exemplification of the present invention.

XX Sequence 9 AA;

Query Match 52.4%; Score 44; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FTVSGNLT 14  
 ID Y06036 standard; Peptide; 9 AA.  
 XX Y06030;  
 AC Y06030;  
 DB 1 fvsngnlt 9

RESULT 8

ID Y06053 standard; Peptide; 9 AA.  
 AC Y06053;  
 XX Y06053;

DT 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 peptide ES09-128.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; cytotoxic T lymphocyte; CTL.

XX Homo sapiens.

OS

XX W09918206-A2.

PN 15-APR-1999.

PD 21-SEP-1998; 98WO-US19609.

PF 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Rosenber SA, Wang RF;

PI WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 45; 88pp; English.

XX Peptide ES09-128 corresponds to amino acid residues 128-136 of

CC human NY ESO-1/CAG-3 ORF1 (see Y05965), a new and potent tumour

CC antigen capable of eliciting an antigen specific immune response

CC by T cells. It was examined for reactivity to a cytotoxic T

CC lymphocyte (CTL), measured as release of granulocyte macrophage

CC colony stimulating factor. Cancer peptides (see Y05967-87) derived

CC from CAG-3, portions of CAG-3 and their variants, are useful as

CC cancer vaccines. A claimed method of preventing or inhibiting

CC cancer involves administering a cancer peptide, with or without an

CC HLA molecule. The cancer peptides form part of, or are derived

CC from, cancers such as primary or metastatic melanoma, thymoma,

CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine

CC cancer, cervical cancer, bladder cancer, kidney cancer and

CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and

CC thyroid cancers.

XX Sequence 9 AA;

Query Match 50.0%; Score 42; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VSGNLTIR 16  
 ID Y06036 standard; Peptide; 9 AA.

Db 1 vsgnltir 9

RESULT 9

ID Y06030 standard; Peptide; 9 AA.

XX Y06030;

AC Y06030;

DB 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;

KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;

KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;

KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;

KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;

KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;

KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

OS W09918206-A2.

PN 15-APR-1999.

PD 21-SEP-1998; 98WO-US19609.

PF 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Rosenber SA, Wang RF;

PI WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a

CC screen for epitopes from the coding region of human ESO-1/CAG-3

CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.

CC The present peptide (ranked 13) corresponds to amino acid residues

CC 127-135 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent

CC tumour antigen capable of eliciting an antigen specific immune

CC response by T cells. Cancer peptides (see Y05967-87) derived from

CC CAG-3, portions of CAG-3 and their variants, are useful as cancer

CC vaccines. A claimed method of preventing or inhibiting cancer

CC involves administering a cancer peptide, with or without an HLA

CC molecule. The cancer peptides form part of, or are derived

CC from, cancers such as primary or metastatic melanoma, thymoma,

CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine

CC cancer, cervical cancer, bladder cancer, kidney cancer and

CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and

CC thyroid cancers.

XX Sequence 9 AA;

Query Match 50.0%; Score 42; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TVSGNLTIR 15

DB 1 tvsgnltir 9

RESULT 10

ID Y06036 standard; Peptide; 9 AA.

```

XX AC Y06036;
XX DT 16-AUG-1999 (first entry)
XX DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.
XX KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;
XX KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
XX KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
XX KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;
XX KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
XX KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
XX KW vaccine; human leukocyte antigen; HLA.
XX OS Homo sapiens.
XX PN W09918206-A2.
XX PD 15-APR-1999.
XX PF 21-SEP-1998; 98WO-US19609.
XX PR 08-OCT-1997; 97US-0061428.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Rosenber SA, Wang RF;
XX DR WPI; 1999-277270/23.
XX PT Cancer antigen NY ESO1/CAG-3
XX PS Example 10; Page 43; 88pp; English.
XX CC This peptide was identified as an HLA peptide motif following a
XX CC screen for epitopes from the coding region of human ESO-1/CAG-3
XX CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.
XX CC The present peptide (ranked 19) corresponds to amino acid residues
XX CC 128-136 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent
XX CC tumour antigen capable of eliciting an antigen specific immune
XX CC response by T cells. Cancer peptides (see Y05967-87) derived from
XX CC CAG-3, portions of CAG-3 and their variants, are useful as cancer
XX CC vaccines. A claimed method of preventing or inhibiting cancer
XX CC involves administering a cancer peptide, with or without an HLA
XX CC molecule. The cancer peptides form part of, or are derived
XX CC from, cancers such as primary or metastatic melanoma, thymoma,
XX CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine
XX CC cancer, cervical cancer, bladder cancer, kidney cancer and
XX CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and
XX CC thyroid cancers.
XX SQ Sequence 9 AA;

Query Match 50.0%; Score 42; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VSGNLTIR 16
Db 1 vsgnltir.9
|||||

RESULT 11
Y79758
ID Y79758 standard; Peptide; 9 AA.
XX AC Y79758;
XX AC
XX AC
XX AC
XX DT 10-MAY-2000 (first entry)
XX DE NY-ESO-1 derived peptide #14.

```

```

XX KW Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;
XX KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;
XX KW cytostatic; melanoma; synovial sarcoma.
XX OS Homo sapiens.
XX PN W0200000824-A1.
XX PD 06-JAN-2000.
XX PF 25-JUN-1999; 99WO-US14493.
XX PR 26-JUN-1998; 98US-0105839.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;
XX PI Chen Y, Gure A, Old LJ;
XX DR WPI; 2000-170933/15.
XX PT Determining the possible presence of breast, endometrial, colorectal,
XX PT lung, bladder or head-neck cancer.
XX PS Example 13; Page 26; 40pp; English.
XX CC A method has been developed for determining the possible presence of a
XX CC cancer, which is not melanoma or synovial sarcoma. The method comprises
XX CC assaying a sample taken from the subject to determine the expression of
XX CC an SSX gene, and determining the expression as a determination of the
XX CC possible presence of cancer. Expression of SSX1 gene indicates possible
XX CC presence of breast, endometrial, colorectal, lung, bladder or head-neck
XX CC cancer. These cancers are also detected by SSX2 and SSX4 gene expression.
XX CC SSX2 gene expression additionally indicates possible presence of
XX CC lymphoma, renal cell cancer, glioma and prostate cancer. Expression of
XX CC SSX4 gene also indicates possible presence of ovarian or stomach cancer.
XX CC SSX5 gene expression indicates the same cancers as SSX1, except breast
XX CC cancer. Determining expression of SSX gene can be used to monitor
XX CC progress of melanoma or synovial sarcoma, which is not cancer. The
XX CC SSX-derived peptide complex stimulates proliferation of cytolytic T
XX CC cells. This is useful for treating cancer, especially melanoma. Y78464
XX CC to Y78468 represent specifically claimed HLA binding peptides for use in
XX CC the method of the invention. 288452 to 288465 represent PCR primers used
XX CC in the isolation of SSX genes in the exemplification of the present
XX CC invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides
XX CC derived from SSX proteins or NY-ESO-1, which are used in the
XX CC exemplification of the present invention.
XX SQ Sequence 9 AA;

Query Match 50.0%; Score 42; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TVSGNLTIT 15
Db 1 tvsgnltit.9
|||||

RESULT 12
Y06027
ID Y06027 standard; Peptide; 9 AA.
XX AC Y06027;
XX AC
XX DT 16-AUG-1999 (first entry)
XX DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.
XX KW NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;
XX KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;

```

KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

OS WO9918206-A2.

XX 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Rosenberg SA, Wang RF;

XX WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 10) corresponds to amino acid residues  
 CC 120-128 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 9 AA;

Query Match 44.0%; Score 37; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLKFTV 8

Db 2 vllkftv 9

RESULT 13

Y06028

ID Y06028 standard; Peptide; 9 AA.

XX Y06028;

XX 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX

OS Homo sapiens.

XX WO9918206-A2.

XX 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Rosenberg SA, Wang RF;

XX WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 11) corresponds to amino acid residues  
 CC 131-139 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 9 AA;

Query Match 44.0%; Score 37; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 NLTIRLT 18

Db 1 nltirlt 8

RESULT 14

Y06000

ID Y06000 standard; Peptide; 10 AA.

XX Y06000;

XX 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX Homo sapiens.

XX WO9918206-A2.

XX 15-APR-1999.

XX

```

PF 21-SEP-1998; 98WO-US19609.
XX
XX
XX 08-OCT-1997; 97US-0061428.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Rosenberg SA, Wang RF;
XX
XX WPI; 1999-277270/23.
XX
XX Cancer antigen NY ESO1/CAG-3
XX
XX Example 10; Page 42; 88pp; English.
XX
XX This peptide was identified as an HLA peptide motif following a
XX screen for epitopes from the coding region of human NY ESO-1/CAG-3
XX ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.
XX The present peptide (ranked 13) corresponds to amino acid residues
XX 131-140 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent
XX tumour antigen capable of eliciting an antigen specific immune
XX response by T cells. Cancer peptides (see Y05967-87) derived from
XX CAG-3, portions of CAG-3 and their variants, are useful as cancer
XX vaccines. A claimed method of preventing or inhibiting cancer
XX involves administering a cancer peptide, with or without an HLA
XX molecule. The cancer peptides form part of, or are derived
XX from, cancers such as primary or metastatic melanoma, thymoma,
XX lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine
XX cancer, cervical cancer, bladder cancer, kidney cancer and
XX adenocarcinomas such as breast, prostate, ovarian, pancreatic and
XX thyroid cancers.
XX
XX Sequence 10 AA;
SQ
Query Match 44.0%; Score 37; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 NLTIRLT 18
Db 1 nltirlt 8

RESULT 15
W39953
ID W39953 standard; peptide; 25 AA.
XX
XX AC W39953;
XX
XX DT 18-JUN-1998 (first entry)
XX
XX Peptide effecting G-protein-coupled receptor activity.
XX
XX G-protein-coupled receptor; GPCR; transmembrane domain; oligomerisation;
XX therapeutic composition; GPCR function; receptor monomeric form;
XX multimeric form; inhibition; GPCR-mediated process; GPCR binding;
XX treatment; disease; alpha adrenergic receptor.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX PN W09800538-A2.
XX
XX PD 08-JAN-1998.
XX
XX PF 01-JUL-1997; 97WO-IB00814.
XX
XX PR 01-JUL-1996; 96US-0021031.
XX
XX (BIOS-) BIOSIGNAL INC.
XX (UYMO-) UNIV MONTREAL.
XX
XX Bouvier M, Dennis M, Hebert TE;

```

```

XX WPI; 1998-086964/08.
XX
XX Peptide(s) or peptide leads affecting G protein-coupled receptor
XX activity - by altering receptor oligomerisation, useful in, e.g.
XX selectively modulating receptor function and treating neurological
XX or genetic diseases
XX
XX Claim 20; Page 59; 75pp; English.
XX
XX Peptides W39953-59 are modelled on transmembrane domains one to seven,
XX respectively, of human alpha adrenergic receptor subtype 2C, a
XX G-protein-coupled receptor (GPCR) whose activity is affected by the
XX formation of oligomers. GPCRs have a recurring pattern unique for the
XX transmembrane domains. The peptides are characterised by the ability
XX to selectively affect oligomerisation of the GPCR from which it was
XX designed. The peptides may be used with suitable carriers in the
XX preparation of therapeutic compositions. They may be administered to
XX selectively modulate GPCR function by affecting the ratio of receptor
XX monomeric to multimeric forms. The compositions can be administered to
XX inhibit GPCR-mediated processes by modulating GPCR binding to its agonist
XX or antagonist. The peptides may also be used to prevent or treat diseases
XX involving GPCRs. They are also useful for in vitro and in vivo studies of
XX GPCRs.
XX
XX Sequence 25 AA;
SQ
Query Match 41.7%; Score 35; DB 19; Length 25;
Best Local Similarity 60.0%; Pred. No. 7 6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 6 FTVSGNLTIT 15
Db 13 ftvsgnltitv 22

Search completed: May 8, 2001, 15:07:08
Job time: 224 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:05:11 ; Search time 62.11 Seconds  
(without alignments)  
5.567 Million cell updates/sec

Title: US-09-165-546A-9

Perfect score: 84

Sequence: 1 VLLKEFTVSGNLTIRLT 18

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 110741

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	39.3	25	1 US-08-087-772A-8	Sequence 8, Appl
2	28	33.3	12	1 US-07-778-233B-13	Sequence 13, Appl
3	28	33.3	12	1 US-07-963-321-13	Sequence 13, Appl
4	28	33.3	12	1 US-08-290-641-13	Sequence 13, Appl
5	28	33.3	12	1 US-08-548-540-13	Sequence 13, Appl
6	28	33.3	12	5 PCT-US96-09809-13	Sequence 13, Appl
7	28	33.3	17	5 5185441-3	Patent No. 5185441
8	27	32.1	21	1 US-08-704-170-28	Sequence 28, Appl
9	27	32.1	21	5 PCT-US94-02631-28	Sequence 28, Appl
10	26	31.0	14	2 US-08-764-640-92	Sequence 92, Appl
11	26	31.0	14	3 US-08-973-225-92	Sequence 92, Appl
12	26	31.0	14	3 US-08-973-225-217	Sequence 217, Appl
13	26	31.0	14	4 US-09-244-298A-92	Sequence 92, Appl
14	26	31.0	18	2 US-08-702-105A-15	Sequence 15, Appl
15	26	31.0	18	3 US-08-702-110A-15	Sequence 15, Appl
16	25	29.8	9	4 US-09-258-754-253	Sequence 253, Appl
17	25	29.8	11	1 US-08-665-966-4	Sequence 4, Appl
18	25	29.8	11	3 US-09-041-780-4	Sequence 4, Appl
19	25	29.8	23	1 US-08-248-505-2	Sequence 2, Appl
20	25	29.8	24	1 US-08-118-270-89	Sequence 89, Appl
21	25	29.8	24	2 US-08-789-078-18	Sequence 18, Appl
22	25	29.8	24	2 US-08-789-078-19	Sequence 19, Appl
23	25	29.8	24	2 US-08-752-633-18	Sequence 18, Appl
24	25	29.8	24	2 US-08-752-633-19	Sequence 19, Appl
25	25	29.8	24	3 US-08-844-978-15	Sequence 15, Appl
26	25	29.8	24	5 PCT-US93-08528-89	Sequence 89, Appl
27	25	29.8	24	5 PCT-US95-04886-18	Sequence 18, Appl

28	25	29.8	24	5 PCT-US95-04886-19	Sequence 19, Appl
29	24	28.6	16	1 US-08-036-555B-121	Sequence 121, App
30	24	28.6	16	1 US-08-469-569-121	Sequence 121, App
31	24	28.6	16	1 US-08-249-322A-121	Sequence 121, App
32	24	28.6	16	1 US-08-469-526A-121	Sequence 121, App
33	24	28.6	16	2 US-08-734-591A-121	Sequence 121, App
34	24	28.6	16	2 US-08-469-660-121	Sequence 121, App
35	24	28.6	16	4 US-08-470-335-121	Sequence 121, App
36	24	28.6	16	4 US-08-735-021-121	Sequence 121, App
37	24	28.6	16	4 US-08-734-664A-121	Sequence 121, App
38	24	28.6	16	5 PCT-US94-05083C-117	Sequence 117, App
39	24	28.6	16	5 PCT-US95-06846A-121	Sequence 121, App
40	24	28.6	18	3 US-09-100-414B-24	Sequence 24, Appl
41	24	28.6	20	1 US-08-440-861-39	Sequence 39, Appl
42	24	28.6	21	1 US-08-290-448A-40	Sequence 40, Appl
43	24	28.6	21	1 US-08-290-448A-51	Sequence 51, Appl
44	24	28.6	21	1 US-08-290-448A-40	Sequence 40, Appl
45	24	28.6	21	1 US-08-290-448A-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1  
US-08-087-772A-8  
; Sequence 8, Application US/08087772A  
; Patent No. 5691155  
; GENERAL INFORMATION:  
; APPLICANT: Namias, Clara  
; APPLICANT: Emorine, Jean L.  
; APPLICANT: Strosberg, Donny A.  
; TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine  
; TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bell, Seitzer, Park & Gibson  
; STREET: Post Office Drawer 34009  
; CITY: Charlotte  
; STATE: No. 5691155th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/087,772A  
; APPLICATION NUMBER: US/08/087,772A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Linker, Raymond O.  
; REGISTRATION NUMBER: 26,419  
; REFERENCE/DOCKET NUMBER: 3339-195  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-087-772A-8

Query Match 39.3%; Score 33; DB 1; Length 25;  
Best Local Similarity 57.1%; Pred. No. 19;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 2 LLKKEFTVSGNLTIRLT 15  
|| || || || ||

Db 8 LLALATVGGNLLVI 21

## RESULT 2

US-07-778-233B-13  
; Sequence 13, Application US/07778233B  
; Patent No. 5270170  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/778,233B  
; FILING DATE: 19911016  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11509-50  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 57 3 0.9  
US-07-778-233B-13

Query Match 33.3%; Score 28; DB 1; Length 12;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLKEFTVSG 10  
| :||| |||  
Db 3 LRREFKVS 11

## RESULT 3

US-07-963-321-13  
; Sequence 13, Application US/07963321  
; Patent No. 5338665  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, Willem P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California

.. 5

; COUNTRY: USA  
; ZIP: 94105

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/963,321  
; FILING DATE: 19921015  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/778,223  
; FILING DATE: 16-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11509-50-1  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 57 3 0.9  
US-07-963-321-13

Query Match 33.3%; Score 28; DB 1; Length 12;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLKEFTVSG 10  
| :||| |||  
Db 3 LRREFKVS 11

## RESULT 4

US-08-290-641-13  
; Sequence 13, Application US/08290641  
; Patent No. 5498530  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, Willem P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/290,641  
; FILING DATE: 15-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,321  
; FILING DATE: 15-OCT-1992  
; APPLICATION NUMBER: US 07/778,223

;  
; FILING DATE: 16-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 11509-50-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 57 3 0.9  
US-08-290-641-13

Query Match 33.3%; Score 28; DB 1; Length 12;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLKEFTVSG 10  
| : || |||  
Db 3 LRREFKVS 11

## RESULT 5

US-08-548-540-13  
; Sequence 13, Application US/08548540  
; Patent No. 5733731  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, Willem P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 162  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/548,540  
; FILING DATE: 26-OCT-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/290,641  
; FILING DATE: 15-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,321  
; FILING DATE: 15-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 16528J-001240US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids

;  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 57 3 0.9  
US-08-548-540-13

Query Match 33.3%; Score 28; DB 1; Length 12;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLKEFTVSG 10  
| : || |||  
Db 3 LRREFKVS 11

## RESULT 6

PCT-US96-09809-13  
; Sequence 13, Application PC/TUS9609809  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; APPLICANT: Cull, Millard G.  
; APPLICANT: Miller, Jeff F.  
; APPLICANT: Stemmer, Willem P.C.  
; TITLE OF INVENTION: Peptide Library and Screening Method  
; NUMBER OF SEQUENCES: 162  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: William M. Smith  
; STREET: One Market Plaza, Stewart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09809  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/548,540  
; FILING DATE: 26-OCT-1995  
; APPLICATION NUMBER: US 08/290,641  
; FILING DATE: 15-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,321  
; FILING DATE: 15-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 16528J-001240US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; CLONE: 57 3 0.9  
PCT-US96-09809-13

Query Match 33.3%; Score 28; DB 5; Length 12;

Best Local Similarity 66.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LLKEFTVSG 10  
| :|||  
Db 3 LRREFKVS 11

## RESULT 7

5185441-3  
; PATENT NO. 5185441  
; APPLICANT: WALLNER, BARBARA P.; HESSESSONS, CATHERINE  
; TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA  
; MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYMPHOCYTE  
; FUNCTION ASSOCIATED ANTIGEN-3  
; NUMBER OF SEQUENCES: 41  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/237,309  
; FILING DATE: 26-AUG-1988  
; SEQ ID NO.3:  
; LENGTH: 17  
5185441-3

Query Match 33.3%; Score 28; DB 6; Length 17;  
Best Local Similarity 50.0%; Pred. No. 88;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TVSGNLTIRLT 18  
||||: ||  
Db 5 TVSGSLTIYLT 16

## RESULT 8

US-08-704-170-28  
; Sequence 28, Application US/08704170  
; Patent No. 5707626  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 No. 5707626th Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,170  
; FILING DATE:  
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/029,850  
; FILING DATE: 11-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-331  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-704-170-28

Query Match 32.1%; Score 27; DB 1; Length 21;  
Best Local Similarity 43.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 LLKEFTVSGNLTIRLT 17  
| :|||  
Db 3 LRREFEYGPRIKRIHM 18

## RESULT 9

PCT-US94-02631-28  
; Sequence 28, Application PC/TUS9402631  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 North Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/02631  
; FILING DATE:  
; CLASSIFICATION:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/029,850  
; FILING DATE: 11-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-331  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003

; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US94-02631-28

Query Match 32.1%; Score 27; DB 5; Length 21;  
Best Local Similarity 43.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 LLKEFTVSGNLTIRLT 17  
| :|||  
Db 3 LRREFEYGPRIKRIHM 18

## RESULT 10

US-08-764-640-92  
; Sequence 92, Application US/08764640  
; Patent No. 5869451

Patent No. 5869451 5837683  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwiria, Steven E.  
APPLICANT: Gates, Christian  
APPLICANT: Schatz, Peter J.  
APPLICANT: Balasubramanian, Palaniappan  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Hendren, Richard W.  
APPLICANT: Deprince, Randolph B.  
APPLICANT: Podduturi, Surekha  
APPLICANT: Yin, Qun  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,640  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 92:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
MOLECULE TYPE: peptide  
US-08-764-640-92

Query Match 31.0%; Score 26; DB 2; Length 14;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LKEFTVSG 10  
|||||  
Db 3 LKEFLHSG 10

RESULT 11  
US-08-973-225-92  
Sequence 92, Application US/08973225A  
Patent No. 6083913  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwiria, Steven E.  
APPLICANT: Duffin, David J.  
APPLICANT: Gates, Christian  
APPLICANT: Haselden, Sherril S.  
APPLICANT: Mattheakis, Larry C.  
APPLICANT: Schatz, Peter J.  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Wrighton, Nicholas C.  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A

THROMBOPOIETIN RECEPTOR  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,225A  
FILING DATE: 04-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3065USW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 92:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 92:  
US-08-973-225-92

Query Match 31.0%; Score 26; DB 3; Length 14;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 LKEFTVSG 10  
|||||  
Db 3 LKEFLHSG 10

RESULT 12  
US-08-973-225-217  
Sequence 217, Application US/08973225A  
Patent No. 6083913  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwiria, Steven E.  
APPLICANT: Duffin, David J.  
APPLICANT: Gates, Christian  
APPLICANT: Haselden, Sherril S.  
APPLICANT: Mattheakis, Larry C.  
APPLICANT: Schatz, Peter J.  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Wrighton, Nicholas C.  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
THROMBOPOIETIN RECEPTOR  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/973,225A  
;; FILING DATE: 04-DEC-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hrubiec, Robert T.  
;; REGISTRATION NUMBER: 36,392  
;; REFERENCE/DOCKET NUMBER: PK3065USW  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 919-248-1000  
;; INFORMATION FOR SEQ ID NO: 217:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: <Unknown>  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-08-973-225-217

Query Match 31.0%; Score 26; DB 3; Length 14;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LKEFTVSG 10  
|||||  
Db 3 LKEFLHSG 10

RESULT 13  
US-09-244-298A-92  
;; Sequence 92, Application US/09244298A  
;; Patent No. 6121238  
;; GENERAL INFORMATION:  
;; APPLICANT: Dower, William J.  
;; APPLICANT: Barrett, Ronald W.  
;; APPLICANT: Cwirla, Steven E.  
;; APPLICANT: Gates, Christian  
;; APPLICANT: Schatz, Peter J.  
;; APPLICANT: Balasubramanian, Palaniappan  
;; APPLICANT: Wagstrom, Christopher R.  
;; APPLICANT: Hendren, Richard W.  
;; APPLICANT: Deprience, Randolph B.  
;; APPLICANT: Podduturi, Surekha  
;; APPLICANT: Yin, Qun  
;; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
;; TITLE OF INVENTION: RECEPTOR  
;; NUMBER OF SEQUENCES: 244  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Glaxo Wellcome  
;; STREET: Five Moore Drive, P.O. Box 13398  
;; CITY: Research Triangle Park  
;; STATE: NC  
;; COUNTRY: USA  
;; ZIP: 27709  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/244,298A  
;; FILING DATE: 11-DEC-1996  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hrubiec, Robert T.  
;; REGISTRATION NUMBER: 36,392  
;; REFERENCE/DOCKET NUMBER: PK3281  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 919-248-1000  
;; INFORMATION FOR SEQ ID NO: 92:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14 amino acids

;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-244-298A-92

Query Match 31.0%; Score 26; DB 4; Length 14;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LKEFTVSG 10  
|||||  
Db 3 LKEFLHSG 10

RESULT 14  
US-08-702-105A-15  
;; Sequence 15, Application US/08702105A  
;; Patent No. 5908839  
;; GENERAL INFORMATION:  
;; APPLICANT: Levitt, Roy C.  
;; APPLICANT: Maloy, W. Lee  
;; APPLICANT: Kari, U. Prasad  
;; APPLICANT: Nicolaides, Nicholas C.  
;; TITLE OF INVENTION: Asthma Associated Factors As Targets For  
;; TITLE OF INVENTION: Treating Allergic Disorders  
;; TITLE OF INVENTION: Disorders  
;; NUMBER OF SEQUENCES: 41  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
;; ADDRESSEE: Dunner L.L.P.  
;; STREET: 1300 I Street N.W., Suite 700  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20005-3315  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/702,105A  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/874,503  
;; FILING DATE: 13-JUN-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fordis, Jean B.  
;; REGISTRATION NUMBER: 32984  
;; REFERENCE/DOCKET NUMBER: 05387.0056-01000  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 408-4000  
;; TELEFAX: (202) 408-4400  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-702-105A-15

Query Match 31.0%; Score 26; DB 2; Length 18;  
Best Local Similarity 62.5%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 TVSGNLT 14  
| : | | |  
Db 7 TTAGNALT 14

```

RESULT 15
US-08-702-110A-15
; Sequence 15, Application US/08702110A
; Patent No. 6037149
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For
; TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related
; TITLE OF INVENTION: Disorders
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,110A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/874,503
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32984
; REFERENCE/DOCKET NUMBER: 05387.0056-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-702-110A-15

```

```

Query Match 31.0%; Score 26; DB 3; Length 18;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 7 TVSGNILT 14
   | : | | |
Db 7 TTAGNALT 14

```

```

Search completed: May 8, 2001, 15:05:12
Job time: 108 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:08:19 ; Search time 68.24 Seconds  
(without alignments)  
18.127 Million cell updates/sec

Title: US-09-165-546A-9

Perfect score: 84

Sequence: 1 VLLKEFTVSGNLTIRLT 18

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 4929

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR\_67:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	39.3	20	2 S45637	oxidoreductase - P
2	29	34.5	12	2 PNO170	alcohol dehydrogen
3	28	33.3	21	2 S58431	phosphatidylinosit
4	26	31.0	17	4 I76673	hypothetical COII/
5	26	31.0	19	2 S74114	3-hydroxyacyl-CoA
6	25	29.8	11	2 I65231	CCK-B gastrin rece
7	25	29.8	20	2 S57286	translation elonga
8	25	29.8	22	2 P00697	hemagglutinin [imp
9	25	29.8	23	2 T44418	carbamoyl-phosphat
10	25	29.8	24	2 A24417	interphotoreceptor
11	24.5	29.2	15	2 I46909	voltage-dependent
12	24	28.6	20	2 S29817	cytochrome P450 2C
13	24	28.6	22	2 B50771	CDw40 antigen Hu54
14	24	28.6	24	2 S38729	probable malate ca
15	23.5	28.0	22	2 C26627	apolipoprotein C-I
16	23	27.4	10	2 PNO165	triose-phosphate i
17	23	27.4	16	2 S36876	aquacobalamin redu
18	23	27.4	16	2 S69361	carbamoyl-phosphat
19	23	27.4	17	2 S58129	hypothetical prote
20	23	27.4	19	2 A47689	flagellar sheath p
21	23	27.4	20	2 PLO039	outer membrane pro
22	23	27.4	21	2 S47200	T-cell receptor J-
23	23	27.4	21	2 S47197	T-cell receptor J-
24	23	27.4	22	2 S47199	T-cell receptor J-
25	23	27.4	24	2 S23121	1H-3-hydroxy-4-oxo
26	23	27.4	25	2 D47689	flagellar core pro
27	22	26.2	12	2 S17540	agglutinin-III lec
28	22	26.2	15	2 PA0097	starch phosphoryla
29	22	26.2	19	2 S34459	ubiquitin - Thermo

30	22	26.2	20	1 LFBSTWT	tet leader peptide
31	22	26.2	20	1 LFBSTU	tet leader peptide
32	22	26.2	20	2 S23742	tet leader peptide
33	22	26.2	21	2 A42762	multicatalytic end
34	22	26.2	22	2 S09021	carboxylesterase (
35	22	26.2	22	2 A33816	Ca2+-transporting
36	21.5	25.6	12	2 PC4377	telomeric and tetr
37	21	25.0	10	2 A27617	triose-phosphate i
38	21	25.0	14	2 B44854	L-2,4-diaminobuty
39	21	25.0	14	2 B39111	Ig heavy chain V r
40	21	25.0	15	2 PA0063	ubiquitin - fungus
41	21	25.0	16	2 I46275	hemoglobin beta x
42	21	25.0	16	2 A61530	triose-phosphate i
43	21	25.0	18	2 PNO175	glutathione transf
44	21	25.0	20	2 S29100	glutathione transf
45	21	25.0	20	2 A60801	acrosome stabilizi

ALIGNMENTS

RESULT 1

S45637

oxidoreductase - Proteus vulgaris (fragment)

C:Species: Proteus vulgaris

C:Date: 10-Dec-1994 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997

C:Accession: S45637

R:Trautwein, T.; Krauss, F.; Lottspeich, F.; Simon, H.

Eur. J. Biochem. 222, 1025-1032, 1994

A:Title: The (2R)-hydroxycarboxylate-viologen-oxidoreductase from Proteus vulgaris is

A:Reference number: S45637; MUID:94298804

A:Accession: S45637

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <TRA>

Query Match 39.3%; Score 33; DB 2; Length 20;

Best Local Similarity 70.0%; Pred. No. 26;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 SGNLTIRLT 18

Db :|||||

6 TGNLTIRLT 15

RESULT 2

PNO170

alcohol dehydrogenase (EC 1.1.1.1) - fungus (Fusarium sporotrichioides) (fragment)

C:Species: Fusarium sporotrichioides

C:Date: 05-Aug-1994 #sequence\_revision 06-Jan-1995 #text\_change 31-Jan-1997

C:Accession: PNO170

R:Fukaya, N.; Chow, L.P.; Suglura, Y.; Tsugita, A.; Ueno, Y.; Tabuchi, K.

submitted to JIPID, May 1994

A:Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotr

A:Reference number: PNO160

A:Accession: PNO170

A:Molecule type: protein

A:Residues: 1-12 <FUK>

A:Experimental source: strain M-1-1

C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 34.5%; Score 29; DB 2; Length 12;

Best Local Similarity 66.7%; Pred. No. 73;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLKEFTVSG 10

Db :|||:||||

2 LLKGITVDG 10

RESULT 3

S58431

phosphatidylinositol transfer protein isoform, 36K - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 26-May-2000  
 C:Accession: S58431  
 R:de Vries, K.J.; Heinrichs, A.A.J.; Cunningham, E.; Brunink, F.; Westerman, J.; Somerha  
 Biochem. J. 310, 643-649, 1995  
 A:Title: An isoform of the phosphatidylinositol-transfer protein transfers sphingomyelin  
 A:Reference number: S58430; MUID:95382786  
 A:Accession: S58431  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-21 <DEV>  
 C:Superfamily: human phosphatidylinositol transfer protein

Query Match 33.3%; Score 28; DB 2; Length 21;  
 Best Local Similarity 75.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLLKEFTV 8  
 I I I I I I  
 Db 1 VLLKEYRV 8

RESULT 4

I76673  
 hypothetical COII/ND5 mutant fusion protein - mouse mitochondrion (fragment)  
 A:Alternate names: COII/ND5 protein  
 C:Species: mitochondrion Mus musculus (house mouse)  
 C:Date: 12-Aug-1996 #sequence\_revision 16-Jul-1998 #text\_change 20-Apr-2000  
 C:Accession: I76673; I76674  
 R:Nelson, I.; Gerasimov, S.; Marsac, C.; Lestienne, P.; Boursot, P.  
 Mamm. Genome 4, 680-683, 1993  
 A:Title: Sequence analysis of a deleted mitochondrial DNA molecule in heteroplasmic mice  
 A:Reference number: I57011; MUID:94108239  
 A:Accession: I76673  
 A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-17 <NEL1>  
 A:Cross-references: GB:S68119; NID:g544777  
 A:Accession: I76674  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 8-17 <NEL2>  
 A:Cross-references: GB:S68119; NID:g544777  
 C:Comment: This is the hypothetical translation of a sequence believed to result from a  
 C:Genetics:  
 A:Gene: COII/ND5  
 A:Genome: mitochondrion  
 A:Genetic code: SGC1  
 C:Keywords: fusion protein; mitochondrion  
 F:1-7/Region: cytochrome-c oxidase chain II  
 F:8-17/Region: NADH dehydrogenase (ubiquinone) chain 5

Query Match 31.0%; Score 26; DB 4; Length 17;  
 Best Local Similarity 45.5%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 LKETVSGNII 13  
 I I I I I I  
 Db 5 LKPLTNNNII 15

RESULT 5

S74114  
 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35) - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 13-Jan-1995 #sequence\_revision 17-Apr-1998 #text\_change 17-Mar-1999  
 C:Accession: S74114  
 R:Dieuaidé-Noubhani, M.; Novikov, D.; Baumgart, E.; Vanhooren, J.C.T.; Franssen, M.; Goet  
 Eur. J. Biochem. 240, 660-666, 1996

A:Title: Further characterization of the peroxisomal 3-hydroxyacyl-CoA dehydrogenases  
 A: acids di- and tri-hydroxycoprostanic acids are metabolized by separate multifunctio  
 A:Reference number: S74113; MUID:97008958  
 A:Accession: S74114  
 A:Molecule type: protein  
 A:Residues: 1-19 <DIE>  
 A:Experimental source: liver  
 A:Superfamily: 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase homol  
 C:Keywords: fatty acid beta-oxidation; mitochondrion; NAD; oxidoreductase

Query Match 31.0%; Score 26; DB 2; Length 19;  
 Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLLKEFTVSG 10  
 I I I I I I  
 Db 4 ILIKHVTVG 13

RESULT 6

I65231  
 CCK-B gastrin receptor isoform - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jul-2000  
 C:Accession: I65231  
 R:Miyaake, A.

Biochem. Biophys. Res. Commun. 208, 230-237, 1995  
 A:Title: A truncated isoform of human CCK-B/gastrin receptor generated by alternative  
 A:Reference number: I52307; MUID:95194412  
 A:Accession: I65231  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-11 <RES>  
 A:Cross-references: GB:S76072; NID:g913752; PIDN:AAB33740.1; PID:g913753  
 C:Genetics:  
 A:Gene: CCK-B

Query Match 29.8%; Score 25; DB 2; Length 11;  
 Best Local Similarity 55.6%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 TVSGNII 15  
 I I I I I I  
 Db 2 SVGGNMLII 10

RESULT 7

S57286  
 translation elongation factor aEF-1 beta - Sulfolobus solfataricus (fragments)  
 C:Species: Sulfolobus solfataricus  
 C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
 C:Accession: S57286  
 R:Arcari, P.; Raimo, G.; Ianniciello, G.; Gallo, M.; Bocchini, V.  
 Biochim. Biophys. Acta 1263, 86-88, 1995  
 A:Title: The first nucleotide sequence of an archaeal elongation factor 1-beta gene.  
 A:Reference number: S57286; MUID:95359209  
 A:Accession: S57286  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <ARC>

Query Match 29.8%; Score 25; DB 2; Length 20;  
 Best Local Similarity 60.0%; Pred. No. 6.1e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLLKEFTVSG 10  
 I I I I I I  
 Db 5 VVLKVPVVG 14

```
RESULT 8
PQ0697
hemagglutinin [imported] : rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: PQ0697
R:Komatsu, S.; Kajiwara, H.; Hirano, H.
Theor. Appl. Genet. 86, 935-942, 1993
A:Title: A rice protein library; a data-file of rice proteins separated by two-dimension
A:Reference number: PQ0696
A:Accession: PQ0697
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <KOM>

Query Match 29.8%; Score 25; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 EFTVSGNI 12
:| | | |
Db 3 KFYVGGNL 10

RESULT 9
T44418
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) small chain [imported]
C:Species: Bacillus stearothermophilus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C:Accession: T44418
R:Vlasikova, H.; Krasny, L.; Fucik, V.; Jonak, J.
submitted to the EMBL Data Library, September 1997
A:Description: The pyrAB gene coding for the large subunit of carbamoylphosphate synthet
A:Reference number: 222760
A:Accession: T44418
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-23 <VLA>
A:Cross-references: EMBL:AJ001805; PIDN:CAA05019.1
A:Experimental source: strain CCM 2184
C:Genetics:
A:Gene: PyrA
A:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain; carbam
C:Keywords: ligase

Query Match 29.8%; Score 25; DB 2; Length 23;
Best Local Similarity 25.0%; Pred. No. 7.1e+02;
Matches 3; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 LLKEFTVSGNIL 13
|:::| |::
Db 10 LIRQFNKGVI 21

RESULT 10
A24417
interphotoreceptor retinoid-binding protein - sheep (fragment)
N:Alternate names: interstitial retinol-binding protein
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Nov-1997
C:Accession: A24417
R:Fong, S.L.; Cook, R.G.; Alvarez, R.A.; Liou, G.I.; Landers, R.A.; Bridges, C.D.B.
FEBS Lett. 205, 309-312, 1986
A:Title: N-terminal sequence homologies in interstitial retinol-binding proteins from 10
A:Reference number: A91365; MUID:86301171
A:Accession: A24417
A:Molecule type: protein
A:Residues: 1-24 <FON>
C:Superfamily: interphotoreceptor retinoid-binding protein
C:Keywords: duplication
```

```
Query Match 29.8%; Score 25; DB 2; Length 24;
Best Local Similarity 38.5%; Pred. No. 7.4e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VLLKEFTVSGNIL 13
| | | | |
Db 12 VLLDNYTFPENLM 24

RESULT 11
I46909
voltage-dependent dihydropyridine-sensitive calcium channel alpha 1 subunit 155 kda i
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 05-Nov-1999
C:Accession: I46909
R:Malouf, N.N.; McMahon, D.K.; Hainsworth, C.N.; Kay, B.K.
Neuron 8, 899-906, 1992
A:Title: A two-motif isoform of the major calcium channel subunit in skeletal muscle.
A:Reference number: I46909; MUID:92265303
A:Accession: I46909
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-15 <MAL>
A:Cross-references: GB:S36895; NID:g249481; PIDN:AAB22180.1; PID:g249482

Query Match 29.2%; Score 24.5; DB 2; Length 15;
Best Local Similarity 41.2%; Pred. No. 5.4e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 1 VLLKEFTVSGNILTIRL 17
| | | | |
Db 2 VLLSLFTI---IFTLEM 15

RESULT 12
S29817
cytochrome P450 2C23 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-Mar-1999
C:Accession: S29817
R:Marie, S.; Rousset, F.; Cresteil, T.
Biochim. Biophys. Acta 1172, 124-130, 1993
A:Title: Age- and tissue-dependent expression of CYP2C23 in the rat.
A:Reference number: S29817; MUID:93176794
A:Accession: S29817
A:Molecule type: mRNA
A:Residues: 1-20 <MAR>
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; transmembrane protein

Query Match 28.6%; Score 24; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 9e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 VSGNILTIRL 17
: | | | | |
Db 7 IIGNLLELNL 16

RESULT 13
B60771
CD440 antigen Hu549 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-May-2000
C:Accession: B60771
R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.
J. Immunol. 142, 562-567, 1989
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-
A:Reference number: A60771; MUID:89093941
A:Accession: B60771
```

Job time: 291 sec

A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-22 <BRA>  
C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 28.6%; Score 24; DB 2; Length 22;  
Best Local Similarity 25.0%; Pred. No. 1.e+03;  
Matches 3; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 4 KEFTVSGNILTI 15  
| : : : |  
Db 9 KQYLINGOCTL 20

RESULT 14  
S38729  
probable malate carrier - Lactococcus lactis (fragment)  
C:Species: Lactococcus lactis  
C:Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 05-Dec-1998  
C:Accession: S38729  
R:Ansanay, V.; Deguin, S.; Blondin, B.; Barre, P.  
FEBS Lett. 332, 74-80, 1993  
A:Title: Cloning, sequence and expression of the gene encoding the malolactic enzyme from Lactococcus lactis  
A:Reference number: S38728; MUID:94009693  
A:Accession: S38729  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-24 <ANS>  
A:Experimental source: strain IL441  
C:Genetics:  
A:Gene: mlep  
C:Function:  
A:Description: catalyzes membrane potential generation via malate/lactate exchange

Query Match 28.6%; Score 24; DB 2; Length 24;  
Best Local Similarity 62.5%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LKEFTVSG 10  
| | | : | |  
Db 4 LKETKISG 11

RESULT 15  
C26627  
apolipoprotein C-I - crab-eating macaque (fragment)  
C:Species: Macaca fascicularis (crab-eating macaque)  
C:Date: 21-May-1998 #sequence\_revision 21-May-1988 #text\_change 31-Dec-1993  
C:Accession: C26627  
R:Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolosi, J.  
Biochemistry 26, 1457-1463, 1987  
A:Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (cynomolgus) monkey  
A:Reference number: A26627; MUID:87185451  
A:Accession: C26627  
A:Molecule type: protein  
A:Residues: 1-22 <HER>  
C:Superfamily: apolipoprotein A-I  
C:Keywords: lipid binding; lipoprotein; VLDL

Query Match 28.0%; Score 23.5; DB 2; Length 22;  
Best Local Similarity 63.6%; Pred. No. 1.2e+03;  
Matches 7; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 3 LKEFTVSGNIL 13  
| | | | |  
Db 11 LKEF---GNTL 18

Search completed: May 8, 2001, 15:08:20



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:06 ; Search time 40.07 seconds  
(without alignments)  
15.388 Million cell updates/sec

Title: US-09-165-546A-9

Perfect score: 84

Sequence: 1 VLLKEFTVSGNLTIRLT 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	32.1	24	1 DHAG_COMTE	P80705 comamonas t
2	25	29.8	22	1 LPL_CORGL	P42456 corynebacte
3	25	28.8	24	1 IRBP_SHEEP	P12863 ovine aries
4	24	28.6	20	1 COG4_CHIOP	P34156 chitonocete
5	23.5	28.0	22	1 APC1_MACEFA	P18657 macaca fasc
6	23	27.4	25	1 FLB3_TREHY	P80161 treponema h
7	22	26.2	20	1 KORA_METTM	P80904 methanobact
8	22	26.2	20	1 LPTB_BACST	P05658 bacillus st
9	22	26.2	20	1 LPTB_BACSU	P23053 bacillus su
10	22	26.2	20	1 TPX_CLOPA	P81361 clostridium
11	22	26.2	23	1 FLA1_SULSH	Q9UW66 sulfolobus
12	21	25.0	10	1 TPIS_NICPL	P19118 nicotiana p
13	21	25.0	12	1 OPS3_DROVI	P17645 drosophila
14	21	25.0	13	1 BP37_LEUMA	P81754 leucophaea
15	21	25.0	15	1 CDN3_LITGI	P56248 litoria gil
16	21	25.0	19	1 MIFH_TRISP	P81529 trichinella
17	21	25.0	20	1 KRC_CLOPA	P80906 methanobact
18	21	25.0	20	1 PGK_CLOPA	P81346 clostridium
19	21	25.0	21	1 DMS_PSECA	P19921 pseudomonas
20	20	23.8	15	1 CDN2_LITGI	P56247 litoria gil
21	20	23.8	15	1 THL_CLOPA	P81347 clostridium
22	20	23.8	16	1 ODPB_SOLTU	P81419 solanus tub
23	20	23.8	17	1 TPIS_PINPS	P81666 pinus pinas
24	20	23.8	19	1 MIFH_TRIMR	P81530 trichuris m
25	20	23.8	20	1 BIP_PHAVU	P80089 phaseolus v
26	20	23.8	20	1 CAP_THICU	P80486 thioabacilli
27	20	23.8	21	1 PSBF_SYNVU	P12239 synechococ
28	20	23.8	22	1 AOFB_MOUSE	Q64133 mus musculu
29	20	23.8	24	1 RS5_VIBPR	P52856 vibrio prot
30	20	23.8	25	1 BGBP_PENVA	P81182 penaeus van
31	19	22.6	13	1 TENC_RANTE	P56918 rana tempor
32	19	22.6	13	1 TEND_RANTE	P56919 rana tempor
33	19	22.6	13	1 TENE_RANTE	P56920 rana tempor

34 19 22.6 17 1 NU4M\_TRIFU Q36834 trichophyto  
35 19 22.6 19 1 FIBA\_MUNMU P14457 muntiacus m  
36 19 22.6 20 1 AROQ\_AMEME P46380 amycolatops  
37 19 22.6 20 1 FRE3\_LITIN P56249 litoria inf  
38 19 22.6 20 1 VR90\_BORPE P81549 bordetella  
39 19 22.6 21 1 EP48\_HUMAN P29322 homo sapien  
40 19 22.6 22 1 HS71\_LEITA P55938 leishmania  
41 19 22.6 23 1 NEUO\_LITCE P81872 litoria cae  
42 19 22.6 24 1 CT31\_LITCI P81851 litoria cit  
43 19 22.6 24 1 LPTN\_ECOLI P09408 escherichia  
44 19 22.6 25 1 IRBP\_PIG P12662 sus scrofa  
45 19 22.6 25 1 YCX8\_ODOSI P49834 odontella s

ALIGNMENTS

RESULT 1  
DHAG\_COMTE STANDARD; PRT; 24 AA.  
AC P80705;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE ALDEHYDE DEHYDROGENASE, GAMMA CHAIN (EC 1.2.99.3) (ALDH) (FRAGMENT).  
OS Comamonas testosteroni (Pseudomonas testosteroni).  
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.  
OX NCBI\_TaxID=285;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=ATCC 15667;  
RA Luykx D.M.A.M., Kim S.W., de Vries S., Duine J.A.;  
RL Submitted (JUL-1996) to the SWISS-PROT data bank.  
CC -1- CATALYTIC ACTIVITY: AN ALDEHYDE + ACCEPTOR + H(2)O = AN ACID +  
CC -1- REDUCED ACCEPTOR.  
CC -1- COFACTOR: MOLYBDENUM.  
CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA  
CHAIN.  
KW Oxidoreductase; Molybdenum.  
FT NON\_TER 24 24  
SQ SEQUENCE 24 AA; 2585 MW; 9E66B518130EA938 CRC64;

Query Match 32.1%; Score 27; DB 1; Length 24;  
Best Local Similarity 45.5%; Pred. No. 1.4e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 EFTVSGNLTIT 15  
:||||:|  
Db 4 QFTVNGRAASI 14

RESULT 2  
LPL\_CORGL STANDARD; PRT; 22 AA.  
ID LPL\_CORGL  
AC P42456;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE LEUA LEADER PEPTIDE.  
GN LEUA.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;  
OC Corynebacterium.  
OX NCBI\_TaxID=1718;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13032;  
RX MEDLINE=94161495; PubMed=8117072;  
RA Patek M., Krumbach K., Eggeling L., Sahn H.;  
RT \*Leucine synthesis in Corynebacterium glutamicum: enzyme activities,  
structure of leuA, and effect of leuA inactivation on lysine

RESULT	4
COG4	CHIOP

```

RESULT 5
APC1_MACFA
ID APC1_MACFA STANDARD: PRT: 22 AA.
AC P18657;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE APOC1-PROTEIN C-I (APO-CI) (FRAGMENT).
GN APOC1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxId=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=87185451; PubMed=3105581;
RA Herbert P.N., Bauserman L.L., Lynch K.M., Saritelli A.L.,
RA Kantor M.A., Nicolosi R.J., Shulman R.S.;
RT "Homologues of the human C and A apolipoproteins in the Macaca
RT fascicularis (Cynomolgus) monkey."
RL Biochemistry 26:1457-1463(1987).
CC -!- FUNCTION: APPEARS TO MODULATE THE INTERACTION OF APOE WITH BETA-
CC MIGRATING VLDL AND INHIBIT BINDING OF BETA-VLDL TO THE LDL
CC RECEPTOR-RELATED PROTEIN.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: SECRETED IN PLASMA.
CC -!- MISCELLANEOUS: APO-CI MAKES UP ABOUT 10% OF THE PROTEIN OF THE
CC VLDL (VERY LOW DENSITY LIPOPROTEIN) & 2% OF THAT OF HDL (HIGH
CC DENSITY LIPOPROTEIN).
CC -!- SIMILARITY: BELONGS TO THE APOC1 FAMILY.

```

```

DR PIR: C26627; C26627.
DR HSP: P02654; 10PP.
KW Plasma: Lipid transport; VLDL.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2349 MW; 4E478A03C91EA914 CRC64;

Query Match 28.0%; Score 23.5; DB 1; Length 22;
Best Local Similarity 63.6%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 3 LKEFTVSGNLT 13
   |||| |||
Db 11 LKEF---GNLT 18

RESULT 6
FLB3 TREHY STANDARD; PRT; 25 AA.
AC P80161;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FLAGELLAR FILAMENT CORE PROTEIN FLAB3 (32 KDA CORE PROTEIN)
DE (FRAGMENT).
GN FLAB3.
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC Bacteria; Spirochaetales; Brachyspiraceae; Brachyspirae.
OX NCBI_TaxID=159;
RN 11
RP SEQUENCE.
RC STRAIN=C5;
RX MEDLINE=93139764; PubMed=1487733;
RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
RA van der Zijst B.A.M., Kusters J.G.;
RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
RT composed of two sheath proteins and three core proteins.";
RL J. Gen. Microbiol. 138:2697-2706(1992).
CC -!- FUNCTION: COMPONENT OF THE CORE OF THE FLAGELLA.
CC -!- SUBUNIT: OUTER LAYER IS COMPOSED OF TWO SHEATH PROTEINS, FLAA1
CC (44 KDA) AND FLAA2 (35 KDA) AND A CORE THAT CONTAINS THREE
CC PROTEINS FLAB1 (37 KDA), FLAB2 (34 KDA) AND FLAB3 (32 KDA).
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC FLAGELLUM.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
DR PIR: D47689; D47689.
KW Flagella; Periplasmic.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2746 MW; 1AA0CA722EE3D0A CRC64;

Query Match 27.4%; Score 23; DB 1; Length 25;
Best Local Similarity 27.3%; Pred. No. 7.3e+02;
Matches 3; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 KEFTVSGNLT 14
   :: :||| :|
Db 14 ROLNLTGNSMT 24

RESULT 7
KORA_METTMM STANDARD; PRT; 20 AA.
AC P80904;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 2-OXOGLUTARATE SYNTHASE SUBUNIT KORA (EC 1.2.7.3) (2-KETOGLUTARATE
DE OXIDOREDUCTASE ALPHA CHAIN) (KOR) (2-OXOGLUTARATE-FERREDOXIN
DE OXIDOREDUCTASE ALPHA SUBUNIT) (FRAGMENT).
GN KORA.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.

```

---

```

OX NCBI_TaxID=79929;
RN 11
RP SEQUENCE.
RX MEDLINE=97261844; PubMed=9108258;
RA Tersteegen A., Linder D., Thauer R.K., Hedderich R.;
RT "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT in Methanobacterium thermoautotrophicum.";
RL Eur. J. Biochem. 244:862-868(1997).
CC -!- CATALYTIC ACTIVITY: 2-OXOBUTANOATE + COA + OXIDIZED FERREDOXIN =
CC PROPANOLYL-COA + CO(2) + REDUCED FERREDOXIN.
CC -!- SUBUNIT: HETEROTETRAMER OF THE KORA, KORB, KORC AND KORD
CC SUBUNITS.
CC -!- MISCELLANEOUS: THE PH OPTIMUM IS PH 9.0 AND THE OPTIMAL
CC TEMPERATURE IS 70 DEGREES CELSIUS.
KW Oxidoreductase.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2080 MW; 112E7E0E4AC76B76 CRC64;

Query Match 26.2%; Score 22; DB 1; Length 20;
Best Local Similarity 37.5%; Pred. No. 8.6e+02;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 KEFTVSGN 11
   :|: :||
Db 2 EEYFIQGN 9

RESULT 8
LPTR_BACST STANDARD; PRT; 20 AA.
AC P05658;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.
GN TETL.
OS Bacillus stearothermophilus, Bacillus cereus, and
OS Staphylococcus hyicus.
OG Plasmid pTHT15, Plasmid pBC16, and Plasmid pSTEL.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1422, 1396, 1284;
RN 11
RP SEQUENCE FROM N.A.
RC PLASMID=pTHT15;
RX MEDLINE=86031344; PubMed=2996983;
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;
RT "Nucleotide sequence of the tetracycline resistance gene of pTHT15, a
RT thermophilic Bacillus plasmid: comparison with staphylococcal TcR
RT controls.";
RL Gene 37:131-138(1985).
RN 12
RP SEQUENCE FROM N.A.
RC PLASMID=pBC16;
RX MEDLINE=90221899; PubMed=21093112;
RA Palva A., Vidgren G., Simonen M., Rintala H., Laamanen P.;
RT "Nucleotide sequence of the tetracycline resistance gene of pBC16
RT from Bacillus cereus.";
RL Nucleic Acids Res. 18:1635-1635(1990).
RN 13
RP SEQUENCE FROM N.A.
RC SPECIES=S.hyicus; PLASMID=pSTEL;
RX MEDLINE=92321725; PubMed=1622166;
RA Schwarz S., Cardoso M., Wegener H.C.;
RT "Nucleotide sequence and phylogeny of the tet(L) tetracycline
RT resistance determinant encoded by plasmid pSTEL from Staphylococcus
RT hyicus.";
RL Antimicrob. Agents Chemother. 36:580-588(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D00006; BAA00004.1; -;  
 EMBL; M11036; AAA22850.1; -;  
 EMBL; X51366; CAA35750.1; -;  
 EMBL; X60828; CAA43219.1; -;  
 PIR; S09233; LFBSTU.  
 PIR; S23742; S23742.  
 Leader peptide; Antibiotic resistance; Plasmid.  
 SEQUENCE 20 AA; 2253 MW; 18D0F4A6CA231CA1 CRC64;

Query Match 26.2%; Score 22; DB 1; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 8.6e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLLKEFTVS 9  
 | | | | : | |  
 Db 9 VOLKEGVS 17

RESULT 9  
 LPTR\_BACSU STANDARD; PRT; 20 AA.

AC P23053;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE TETRACYCLINE RESISTANCE LEADER PEPTIDE.

GN TEPL.  
 OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=1423;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=89000797; PubMed=2844262;

RA Sakaguchi R., Anano H., Shishido K.;

RT "Nucleotide sequence homology of the tetracycline-resistance

RT determinant naturally maintained in Bacillus subtilis Marburg 168

RT chromosome and the tetracycline-resistance gene of B. subtilis

RT plasmid pNS1981.";

RL Biochim. Biophys. Acta 950:441-444(1988).

[2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91257555; PubMed=1646143;

RA Anano H., Sakaguchi R., Shishido K.;

RT "An insertion of Escherichia coli transposable element IS1K into the

RT site immediately before tetracycline-resistance determinant of

RT Bacillus subtilis chromosomal DNA fragment in cloning in E. coli.";

RL FEMS Microbiol. Lett. 63:5-8(1991).

CC

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X08034; CAA30826.1; -;  
 DR EMBL; X58999; CAA41744.1; -;  
 DR EMBL; Z90124; CAB16115.1; -;  
 DR PIR; S04802; LFBSTT.  
 DR PIR; S17282; S17282.

DR Subtilist; BG11051; tetL.

DR Leader peptide; Antibiotic resistance.

KW SEQUENCE 20 AA; 2298 MW; E870F516CA23111A CRC64;

SQ

Query Match 26.2%; Score 22; DB 1; Length 20;  
 Best Local Similarity 66.7%; Pred. No. 8.6e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLLKEFTVS 9  
 | | | | : | |  
 Db 9 VOLKEGVS 17

RESULT 10

TPX\_CLOPA

ID TPX\_CLOPA STANDARD; PRT; 20 AA.

AC P81361;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE PROBABLE THIOL PEROXIDASE (EC 1.11.1.-) (CP 42) (FRAGMENT).

OS Clostridium pasteurianum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1501;

[1]

RP SEQUENCE.

RC STRAIN=W5;

RX MEDLINE=98291870; PubMed=9629918;

RA Flengsrud R., Skjeldal L.;

RT "Two-dimensional gel electrophoresis separation and N-terminal

RT sequence analysis of proteins from Clostridium pasteurianum W5.";

RL Electrophoresis 19:802-806(1998).

CC -!- FUNCTION: HAS ANTI-OXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR

CC H(2)O(2) (BY SIMILARITY).

CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:

CC 4.6, ITS MW IS: 20.2 KDA.

CC -!- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.

DR InterPro; IPR002065; .

DR PROSITE; PS01265; TPX; PARTIAL.

KW Oxidoreductase; Peroxidase.

FT NON\_TER 20

SQ SEQUENCE 20 AA; 2193 MW; 08178FCD2782E765 CRC64;

Query Match 26.2%; Score 22; DB 1; Length 20;  
 Best Local Similarity 40.0%; Pred. No. 8.6e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 TVSGNLTIR 16  
 | | | | : | |  
 Db 4 TFOGNEVTIQ 13

RESULT 11

FLAL\_SULSH

ID FLAL\_SULSH STANDARD; PRT; 23 AA.

AC Q9UWG6;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE 31/33 KDA FLAGELLIN (FRAGMENT).

OS Sulfolobus shibatae.

OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.

OX NCBI\_TaxID=2286;

[1]

RP SEQUENCE.

RC STRAIN=B12.

RX MEDLINE=96146545; PubMed=8550530;

RA Faguy D.M., Bayley D.P., Kostyukova A.S., Thomas N.A., Jarrell K.F.;

RT "Isolation and characterization of flagella and flagellin proteins

RT from the Thermocacidophilic archaea Thermoplasma volcanum and

RT Sulfolobus shibatae.";

RL J. Bacteriol. 178:902-905(1996).

CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO

CC FORM THE FILAMENTS OF FLAGELLA.

CC -!- PTM: GLYCOSYLATED.  
 CC -!- SIMILARITY: BELONGS TO THE ARCHAEL FLAGELLIN FAMILY.  
 KW Flagella; Glycoprotein.  
 FT NON\_TER 23 23  
 SQ SEQUENCE 23 AA; 2362 MW; BB372157B44DB3AC CRC64;

Query Match 26.2%; Score 22; DB 1; Length 23;  
 Best Local Similarity 23.1%; Pred. No. 9.9e+02;  
 Matches 3; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLLKFTVSGNII 13  
 :| | :| :|  
 Db 8 ILLIATITASVL 20

RESULT 12  
 TPIS\_NICPL STANDARD; PRT; 10 AA.  
 AC P19118;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE TRIOSEPHOSPHATE ISOMERASE, CYTOSOLIC (EC 5.3.1.1) (TIM) (FRAGMENT).  
 OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 CC Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4092;  
 RN [1]  
 RP SEQUENCE.

RA Bauw G., de Loose M., Inze D., van Montagu M., Vandekerckhove J.;  
 RT "Alterations in the phenotype of plant cells studied by NH2-terminal  
 RT amino acid-sequence analysis of proteins electrophoretically separated from two-  
 RT dimensional gel-separated total extracts";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4806-4810(1987).  
 CC -!- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE - DIHYDROXY-  
 CC ACETONE PHOSPHATE.  
 CC -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.  
 DR PIR; A27617; A27617.  
 DR InterPro; IPR000652;  
 DR PROSITE; PS00171; TIM; PARTIAL.  
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;  
 KW Pentose shunt.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1140 MW; 80B9D37862C9C9D1 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 6e+02;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 KEFTVSGN 11  
 :| | :| :|  
 Db 2 RTFFVGGN 9

RESULT 13  
 OPS3\_DROVI STANDARD; PRT; 12 AA.  
 AC P17645;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE OPSIN RH3 (INNER R7 PHOTORECEPTOR CELLS OPSIN) (FRAGMENT).  
 GN RH3.  
 OS Drosophila virilis (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7244;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90249748; PubMed=2140105;  
 RA Fortini M.E., Rubin G.M.;  
 RT "Analysis of cis-acting requirements of the Rh3 and Rh4 genes reveals  
 RT a bipartite organization to rhodopsin promoters in Drosophila  
 RT melanogaster";  
 RL Genes Dev. 4:444-463(1990).  
 CC -!- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
 CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
 CC LINKED TO CIS-RETINAL.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- MISCELLANEOUS: EACH DROSOPHILA EYE IS COMPOSED OF 800 FACETS OR  
 CC OMMATIDIA. EACH OMMATIDIUM CONTAINS 8 PHOTORECEPTOR CELLS (R1-R8),  
 CC THE R1 TO R6 CELLS ARE OUTER CELLS, WHILE R7 AND R8 ARE INNER  
 CC CELLS.  
 CC -!- MISCELLANEOUS: OPSIN RH3 IS SENSITIVE TO UV LIGHT.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC OPSIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X51350; CAA35742.1; -  
 DR GCRDB; GCR\_0779; -  
 DR FlyBase; FBgn0013091; Dvir\Rh3.  
 DR InterPro; IPR000276; -  
 DR InterPro; IPR001760; -  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_FL1; PARTIAL.  
 DR PROSITE; PS00238; OPSIN; PARTIAL.  
 KW Photoreceptor; Retinal protein; Transmembrane; Phosphorylation;  
 KW Glycoprotein; G-protein coupled receptor; Vision.  
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1253 MW; 04024E43495865B0 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 12;  
 Best Local Similarity 50.0%; Pred. No. 7.4e+02;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 EFTVSG 10  
 :| :| :|  
 Db 2 DFNISG 7

RESULT 14  
 BP37\_LEUMA STANDARD; PRT; 13 AA.  
 AC P81754;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE BRAIN PROTEIN 37F3.  
 OS Leucophaea maderae (Madeira cockroach).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
 CC Blaberoidea; Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Brain;  
 RX MEDLINE=97269266; PubMed=9114447;  
 RA Muren J.E., Naessel D.R.;  
 RT "Seven tachykinin-related peptides isolated from the brain of the

RT madeira cockroach; evidence for tissue-specific expression of  
isoforms.";  
RL Peptides 18:7-15(1997).  
CC -!- TISSUE SPECIFICITY: BRAIN.  
CC -!- MASS SPECTROMETRY: MW=1435.7; METHOD=MALDI-MS.  
SQ SEQUENCE 13 AA; 1436 MW; 8E532C9DE9A7D1B5 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 13;  
Best Local Similarity 62.5%; Pred. No. 8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Oy 2 LKFTVS 9  
| : | | | |  
Db 2 LFEESTVS 9

RESULT 15  
CDN3\_LITGI STANDARD; PRT; 15 AA.  
AC P56248;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE CAERIDIN 3.  
OS Litoria gilleni.  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
CC Litoria.  
OX NCBI\_TaxID=39405;  
RN [1]  
RP SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE-Parotoid gland;  
RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;  
RT "Peptides from Australian frogs. The structures of the caerins and  
caeridins from Litoria gilleni.";  
RL J. Chem. Res. 139:937-961(1993).  
CC -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR  
ANTIBIOTIC ACTIVITY.  
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
GLANDS.  
CC -!- MASS SPECTROMETRY: MW=1428; METHOD-FAB.  
KW Amphibian skin; Amidation.  
FT MOD.RES 15 15 AMIDATION.  
SQ SEQUENCE 15 AA; 1430 MW; 06E90A797AF70CBF CRC64;

Query Match 25.0%; Score 21; DB 1; Length 15;  
Best Local Similarity 50.0%; Pred. No. 9.3e+02;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Oy 6 FTVSGNIL 13  
| : | | | |  
Db 3 FDAIGNLL 10

Search completed: May 8, 2001, 15:14:07  
Job time: 528 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:16:05 ; Search time 114.89 Seconds  
(without alignments)  
18.363 Million cell updates/sec

Title: US-09-165-546A-9  
Perfect score: 84  
Sequence: 1 VLLKEFTVSGNLTIRLT 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_15.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_unclassified.\*  
13: sp\_vertebrate.\*  
14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	39.3	20	Q9R4Y1	Q9r4y1 proteus vul
2	29	34.5	21	Q9TR36	Q9tr36 bos taurus
3	29	34.5	22	Q9ZAW2	Q9zaw2 francisella
4	29	34.5	24	Q52390	Q52390 nodularia s
5	28	33.3	21	Q9TR37	Q9tr37 bos taurus
6	27.5	32.7	20	Q9TRU5	Q9tru5 oryctolagus
7	27	32.1	15	Q9UWH9	Q9uwh9 thermococcu
8	26	31.0	19	Q9N195	Q9n195 bos taurus
9	26	31.0	24	P92152	P92152 caenorhabdi
10	25	29.8	16	Q47605	Q47605 escherichia
11	25	29.8	21	Q9T167	Q9t167 bacterioph
12	25	29.8	23	O50301	O50301 bacillus st
13	25	29.8	23	Q28016	Q28016 bos taurus
14	25	29.8	13	P82398	P82398 litoria ran
15	25	29.8	24	Q9TY69	Q9ty69 drosophila
16	24.5	28.2	15	Q28822	Q28822 oryctolagus
17	24	28.6	19	Q9QVK0	Q9qvk0 mus sp. mep
18	24	28.6	20	Q9S888	Q9s888 lupinus alb
19	24	28.6	20	Q9QW31	Q9qw31 rattus sp.

20	24	28.6	23	13	P82399	P82399 litoria aur
21	24	28.6	24	4	Q16475	Q16475 homo sapien
22	24	28.6	24	4	Q9UMB2	Q9umb2 homo sapien
23	24	28.6	24	11	O88717	O88717 mus musculus
24	24	28.6	25	14	Q9YN05	Q9yn05 myxoma viru
25	24	28.6	25	14	Q9PXH2	Q9pxh2 human t-cel
26	23	27.4	11	11	Q9QYF6	Q9qyf6 mus musculus
27	23	27.4	16	2	Q52901	Q52901 rhizobium m
28	29	27.4	16	13	Q9PRQ0	Q9prq0 oncorhynch
29	23	27.4	17	7	Q31213	Q31213 mus musculus
30	23	27.4	17	10	O24445	O24445 ceratodon p
31	23	27.4	17	14	Q9IHJ0	Q9ihj0 human polio
32	23	27.4	17	14	Q9IHI9	Q9ihi9 human polio
33	23	27.4	17	14	Q9IHI8	Q9ihi8 human polio
34	23	27.4	17	14	Q9IHI7	Q9ihi7 human polio
35	23	27.4	17	14	Q9IHI6	Q9ihi6 human polio
36	23	27.4	17	14	Q9IHI5	Q9ihi5 human polio
37	23	27.4	17	14	Q9IHI4	Q9ihi4 human polio
38	23	27.4	17	14	Q9IHI3	Q9ihi3 human polio
39	23	27.4	17	14	Q9IHI2	Q9ihi2 human polio
40	23	27.4	17	14	Q9IHI1	Q9ihi1 human polio
41	23	27.4	17	14	Q9IHI0	Q9ihi0 human polio
42	23	27.4	17	14	Q9IHH9	Q9ihh9 human polio
43	23	27.4	17	14	Q9IHH8	Q9ihh8 human polio
44	23	27.4	17	14	Q9IHH7	Q9ihh7 human polio
45	23	27.4	17	14	Q9IHH6	Q9ihh6 human polio

## ALIGNMENTS

RESULT 1  
ID Q9R4Y1 PRELIMINARY; PRT; 20 AA.  
AC Q9R4Y1; PRELIMINARY; PRT; 20 AA.  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE (2R)-HYDROXYCARBOXYLATE-VIOGEN-OXIDOREDUCTASE (FRAGMENT).  
OS Proteus vulgaris.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Proteus.  
OX NCBI\_TaxID=585;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-94298804; PubMed-8026480;  
RA Trautwein T., Krauss F., Lottspeich F., Simon H.;  
RT "The (2R)-hydroxycarboxylate-viologen-oxidoreductase from Proteus  
vulgaris is a molybdenum-containing iron-sulphur protein.";  
RL Eur. J. Biochem. 222:1025-1032(1994).  
SQ SEQUENCE 20 AA; 2145 MW; D12A6CFC51207C63 CRC64;

Query Match 39.3%; Score 33; DB 2; Length 20;  
Best Local Similarity 70.0%; Pred. No. 54;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 SGNLTIRLT 18  
:|||||  
Db 6 TGNLTIRLT 15

RESULT 2  
ID Q9TR36 PRELIMINARY; PRT; 21 AA.  
AC Q9TR36;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE PHOSPHATIDYLINOSITOL-TRANSFER PROTEIN 36 KDA ISOFORM (FRAGMENT).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX MEDLINE=95382786; PubMed=7654206;
RP de Vries K.J., Heinrichs A.A., Cunningham E., Brunink F.,
RA Westernman J., Somerharju P.J., Cockcroft S., Wirtz K.W., Snoek G.T.;
RA "An isoform of the phosphatidylinositol-transfer protein transfers
RT sphingomyelin and is associated with the Golgi system.";
RL Blochem. J. 310:643-649(1995).
DR INTERPRO: IPR001666;
DR PFAM: PF02121; IP_trans; 1.
SQ SEQUENCE 21 AA; 2442 MW; 0978BC2B0067EF0E CRC64;

Query Match 34.5%; Score 29; DB 6; Length 21;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLLKEFTV 8
   |||:|:|
Db 1 VLIKEFRV 8

RESULT 3
Q9ZAW2 PRELIMINARY; PRT; 22 AA.
ID Q9ZAW2
AC Q9ZAW2
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE ENVA (FRAGMENT).
OS ENVA.
GN Francisella tularensis.
OC Bacteria; Proteobacteria; gamma subdivision; Francisella group;
OC Francisella.
OX NCBI_TaxID=263;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LVS (LIVE VACCINE STRAIN);
RA Salimnia H., Clairoux N., Bolssinet M.;
RT "Cloning and characterization of Francisella tularensis LVS homolog of
RT the E. coli cell division gene ftsZ.";
RL Submitted (OCT-1996) to the EMBL/genbank/DBJ databases.
DR EMBL: U76309; AAC99559.1;
FT NON_TER 22
SQ SEQUENCE 22 AA; 2333 MW; ADA9026DE208C0A5 CRC64;

Query Match 34.5%; Score 29; DB 2; Length 22;
Best Local Similarity 55.6%; Pred. No. 3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LLLKEFTVSG 10
   |||:|:|
Db 6 IAKFFSVTG 14

RESULT 4
Q52390 PRELIMINARY; PRT; 24 AA.
ID Q52390
AC Q52390;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SERINE/THREONINE PHOSPHATASE 2B (FRAGMENT).
GN PF2B-CYANO1.
OS Nodularia spumigena.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nodularia.
OX NCBI_TaxID=70799;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L575;

Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
MEDLINE=98047369; PubMed=9385145;
Shi L., Carmichael W.W.;
"ppl-cyano2, a protein serine/threonine phosphatase 1 gene from the
cyanobacterium Microcystis aeruginosa UTEX 2063.";
Arch. Microbiol. 168:528-531(1997).
EMBL: AF037161; AAC46049.1;
INTERPRO: IPR000934;
PFAM: PF00149; Sphosphatase; 1.
NON_TER 1
FT NON_TER 24
SQ SEQUENCE 24 AA; 2768 MW; DB9C36D51C0DA51D CRC64;

Query Match 34.5%; Score 29; DB 2; Length 24;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LLKEFTVSGN 11
   |||:|:|
Db 5 LILKFEVGGN 14

RESULT 5
Q9TR37 PRELIMINARY; PRT; 21 AA.
ID Q9TR37
AC Q9TR37;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PHOSPHATIDYLINOSITOL-TRANSFER PROTEIN 35 KDA ISOFORM (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=95382786; PubMed=7654206;
de Vries K.J., Heinrichs A.A., Cunningham E., Brunink F.,
Westernman J., Somerharju P.J., Cockcroft S., Wirtz K.W., Snoek G.T.;
"An isoform of the phosphatidylinositol-transfer protein transfers
RT sphingomyelin and is associated with the Golgi system.";
RL Blochem. J. 310:643-649(1995).
DR INTERPRO: IPR001666;
DR PFAM: PF02121; IP_trans; 1.
SQ SEQUENCE 21 AA; 2462 MW; ESA50C9DA8C4F2D6 CRC64;

Query Match 33.3%; Score 28; DB 6; Length 21;
Best Local Similarity 75.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLLKEFTV 8
   |||:|:|
Db 1 VLLKEYRV 8

RESULT 6
Q9TRU5 PRELIMINARY; PRT; 20 AA.
ID Q9TRU5
AC Q9TRU5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE FACTOR IX LIGHT CHAIN (GLA DOMAIN).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=92132840; PubMed=1776139;
Rao L.V., Hoang A.D.;
```

Query Match

```
Db 1 :||| ||
3 LLDFTIGN 11

RESULT 11
Q9T167 O9T167 PRELIMINARY; PRT; 21 AA.
AC O9T167;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE GP53.
OS Bacteriophage A118.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
OX NCBI_TaxID=40521;
[1]
RN RP SEQUENCE FROM N.A.
RA Loessner M.J.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96020653; PubMed=8577256;
RA Loessner M.J., Wendlinger G., Scherer S.;
RT "Heterogeneous endolysins in Listeria monocytogenes bacteriophages: a
RT new class of enzymes and evidence for conserved holin genes within the
RT siphoviral lysis cassettes.";
RL Mol. Microbiol. 16:1231-1241(1995).
[3]
RN RP SEQUENCE FROM N.A.
RA Loessner M.J., Inman R.B., Lauer P., Calendar R.;
RT "Complete nucleotide sequence, molecular analysis and genome structure
RT of bacteriophage A118 of Listeria monocytogenes: implications for
RT phage evolution.";
RL Mol. Microbiol. 35:324-340(2000).
DR EMBL; AJ242593; CAB53843.1; -.
DR INTERPRO; IPR000217; -.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG. 1.
SQ SEQUENCE 21 AA; 2485 MW; 7B5D940B75D7CF90 CRC64;

Query Match 29.8%; Score 25; DB 9; Length 21;
Best Local Similarity 40.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LKEFTVSGNI 12
Db 1 MREIEIYGN 10

RESULT 12
O50301 ID O50301 PRELIMINARY; PRT; 23 AA.
AC O50301;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
DE CARBAMOYLPHOSPHATE SYNTHETASE SMALL SUBUNIT (FRAGMENT).
GN PYRAA.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1422;
[1]
RN RP SEQUENCE FROM N.A.
RA STRAIN=CCM 2184.
RA Vlaskova H., Krasny L., Fucik V., Jonak J.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ001805; CAA05019.1; -.
FT NON_TER 1
SQ SEQUENCE 23 AA; 2690 MW; 90464123B3486524 CRC64;

Query Match 29.8%; Score 25; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 LKREFTVSGNII 13
Db 10 LIRQFNKKGEVI 21

RESULT 13
Q28016 ID Q28016 PRELIMINARY; PRT; 23 AA.
AC Q28016;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE INSULIN-LIKE GROWTH FACTOR-1 RECEPTOR (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=96432479; PubMed=8835545;
RA Moody D.E., Pomp D., Barendse W.;
RT "Linkage mapping of the bovine insulin-like growth factor-1 receptor
RT gene.";
RL Mamm. Genome 7:168-169(1996).
DR EMBL; U33122; AAB52601.1; -.
FT NON_TER 1
FT NON_TER 23
SQ SEQUENCE 23 AA; 2458 MW; D9814EC8DB0FE853 CRC64;

Query Match 29.8%; Score 25; DB 6; Length 23;
Best Local Similarity 45.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 VSGNLTIRLT 18
Db 7 ISGNVCPDRST 17

RESULT 14
P82398 ID P82398 PRELIMINARY; PRT; 23 AA.
AC P82398;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE AUREIN 4.2.
OS Litoria raniformis, and Litoria aurea (Australian frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=116057, 8371;
[1]
RN RP SEQUENCE.
RA Rozek T., Wegener K.L., Bowie J.H., Oliver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer aurein peptides from the Australian
RT bell frogs Litoria aurea and Litoria raniformis.";
RL Eur. J. Biochem. 0:0-0(2000).
CC -!- FUNCTION: HAS NO ANTIMICROBIAL OR ANTICANCER ACTIVITY.
SQ SEQUENCE 23 AA; 2403 MW; 2ECE31DF272EAF6 CRC64;

Query Match 29.8%; Score 25; DB 13; Length 23;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 6; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 3 LKREFTVSGNII 14
Db 1 :||| :| :||
```

Db 10 LKEF--AGGVVT 19

RESULT 15

Q9TY69  
ID Q9TY69 PRELIMINARY: PRT: 24 AA.  
AC Q9TY69;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE PROTEIN PHOSPHATASE FROM PCR FRAGMENT D14 (FRAGMENT).  
GN PPD14.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92331787; PubMed=1321058;  
RA Chen M.X., Chen Y.H., Cohen P.T.;  
RT "Polymerase chain reactions using Saccharomyces, Drosophila and human  
RT DNA predict a large family of protein serine/threonine phosphatases.";  
RL FEBS Lett. 306:54-58(1992).  
DR EMBL: S39996; AAB22466.1; -.  
DR HSSP: P08129; 1FJM.  
DR FLYBASE; FBgn005780; Ppd14.  
DR INTERPRO: IPR000934; -.  
DR PFAM: PF00149; STphosphatase; 1.  
FT NON\_TER 1  
SQ SEQUENCE 24 AA; 2716 MW; 9746395FCC1E7FBB CRC64;

Query Match 29.8%; Score 25; DB 5; Length 24;  
Best Local Similarity 46.7%; Pred. No. 1.6e+03;  
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 LKKEFTVSGNLTIR 16  
|:|:|:|:|:  
Db 5 LVKLFVGGPPATR 19

Search completed: May 8, 2001, 15:16:06  
Job time: 532 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:07:08 ; Search time 113.31 Seconds  
(without alignments)  
9.081 Million cell updates/sec

Title: US-09-165-546A-11

Perfect score: 105

Sequence: 1 GAASGLNGCCRGARGPE 18

Scoring table:

BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 165078

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_0401.\*

1: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	105	100.0	18	Y52438	Human tumour antig
2	64	61.0	10	Y06012	Human cancer antig
3	63	60.0	10	Y06063	Human cancer antig
4	63	60.0	10	Y06008	Human cancer antig
5	59	56.2	9	Y06044	Human cancer antig
6	57	54.3	10	Y06060	Human cancer antig
7	57	54.3	10	Y05994	Human cancer antig
8	53	50.5	9	Y06047	Human cancer antig
9	47	44.8	10	Y06059	Human cancer antig
10	47	44.8	10	Y05993	Human cancer antig
11	42	40.0	10	R69313	Gp IIb/IIIa recept

12	42	40.0	10	19	W52088
13	42	40.0	10	21	Y95511
14	41	39.0	11	21	Y97091
15	41	39.0	11	21	Y97109
16	41	39.0	20	21	Y97089
17	41	39.0	20	21	Y97107
18	40	38.1	11	21	Y82639
19	40	38.1	12	20	Y41492
20	40	38.1	25	19	W82534
21	39.5	37.6	15	16	R75282
22	38	36.2	23	10	P90448
23	38	36.2	24	21	B14992
24	37	35.2	15	20	Y33078
25	37	35.2	25	20	W73433
26	37	35.2	25	20	W87739
27	36	34.3	15	18	W38912
28	36	34.3	15	18	W24889
29	36	34.3	15	18	W12744
30	36	34.3	24	21	B51768
31	35	33.3	15	10	P90207
32	35	33.3	24	22	B65047
33	35	33.3	25	11	R07983
34	34.5	32.9	25	21	B21617
35	34	32.4	25	21	B14994
36	33.5	31.9	15	21	B03867
37	33.5	31.9	16	20	Y24163
38	33	31.4	9	13	R26401
39	33	31.4	9	14	R69302
40	33	31.4	9	15	Y08517
41	33	31.4	9	19	W50585
42	33	31.4	9	21	Y95453
43	33	31.4	9	21	Y79751
44	33	31.4	9	21	Y54968
45	33	31.4	10	19	W52077

#### ALIGNMENTS

RESULT 1  
Y52438  
ID Y52438 standard; Protein; 18 AA.  
XX  
AC Y52438;  
XX  
DT 15-FEB-2000 (first entry)  
XX  
DE Human tumour antigen NY-ESO-1 peptide #11.  
XX  
Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;  
KW T-cell; helper; stimulation; proliferation; treatment;  
KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;  
KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;  
KW lymphoma.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO9953938-A1.  
XX  
PD 28-OCT-1999.  
XX  
PF 24-MAR-1999; 99WO-US06875.  
XX  
PR 17-APR-1998; 98US-0062422.  
XX  
PR 02-OCT-1998; 98US-0165546.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
PI Gure A, Ritter G;  
XX  
DR WPI; 2000-038483/03.

Targeting peptide  
GPIIb/IIIa recepto  
Tau conotoxin sequ  
Tau conotoxin sequ  
Tau conotoxin sequ  
Tau conotoxin sequ  
Influenza A virus  
Fragment of human  
HIV-1 p24 epitope  
A-lineage conotoxi  
Sequence of mature  
Partial sequence o  
Carbohydrate antig  
Human secreted pro  
Secreted protein e  
Peptide resembling  
Predatory cone sna  
A-lineage conotoxi  
Human secreted pro  
Antigenic peptide  
Gene #7 associated  
Gag protein T-Cell  
Cone snail alpha-c  
Partial sequence o  
GIF receptor targe  
Alpha-conotoxin pe  
Sequence of techn  
Gp IIb/IIIa recept  
Endothelin metal c  
GPIIb/IIIa recepto  
NY-ESO-1 derived p  
Peptide ligand for  
Targeting peptide

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
 PT useful for therapeutic and diagnostic purposes -  
 PS Claim 4; Page 22; 49pp; English.  
 XX Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels  
 CC in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and  
 CC tumour cell lines with some degree of frequency - these included  
 CC melanoma specimens and cell lines, and breast and bladder cancer  
 CC specimens, with expression in other tumour types being sporadic.  
 CC These NY-ESO-1-derived peptides may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of  
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
 CC or lymphoma) and to stimulate the proliferation of T cells.  
 XX Sequence 18 AA;  
 SQ

Query Match 100.0%; Score 105; DB 21; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAASGLNGCCRCGARGPE 18  
 Db 1 gaasgngccrcgargpe 18  
 ||||||||||||||||

RESULT 2  
 Y06012  
 ID Y06012 standard; Peptide; 10 AA.  
 XX  
 AC Y06012;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.  
 XX  
 KW NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PA Rosenberg SA, Wang RF;  
 XX  
 PI WPI; 1999-277270/23.  
 XX  
 DR Cancer antigen NY ESO1/CAG-3  
 XX  
 PS Example 10; Page 42; 88pp; English.  
 XX

CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 25) corresponds to amino acid residues  
 CC 71-80 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer, and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX Sequence 10 AA;  
 SQ

Query Match 61.0%; Score 64; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 GLNGCCRCGA 14  
 Db 1 glngccrcga 10  
 ||||||||

RESULT 3  
 Y06063  
 ID Y06063 standard; Peptide; 10 AA.  
 XX  
 AC Y06063;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX  
 DE Human cancer antigen NY ESO-1/CAG-3 peptide ESO10-72.  
 XX  
 KW NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; cytotoxic T lymphocyte; CTL.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PA Rosenberg SA, Wang RF;  
 XX  
 PI WPI; 1999-277270/23.  
 XX  
 DR Cancer antigen NY ESO1/CAG-3  
 XX  
 PT Example 10; Page 45; 88pp; English.  
 XX  
 PS Peptide ESO10-72 corresponds to amino acid residues 72-81 of  
 CC human NY ESO-1/CAG-3 ORF1 (see Y05965), a new and potent tumour  
 CC antigen capable of eliciting an antigen specific immune response  
 CC by T cells. It was examined for reactivity to a cytotoxic T  
 CC lymphocyte (CTL), measured as release of granulocyte macrophage  
 CC colony stimulating factor.. Cancer peptides (see Y05967-87) derived

CC from CAG-3, portions of CAG-3 and their variants, are useful as  
 CC cancer vaccines. A claimed method of preventing or inhibiting  
 CC cancer involves administering a cancer peptide, with or without an  
 CC HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX Sequence 10 AA;

Query Match 60.0%; Score 63; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.035;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LNCRCRCGAR 15  
 Db 1 lncrcrcgar 10

RESULT 4  
 Y06008  
 ID Y06008 standard; Peptide; 10 AA.

XX AC Y06008;

DT 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX OS Homo sapiens.  
 XX PN WO9918206-A2.

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX PA (USSR ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenberg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX PT Cancer antigen NY ESO1/CAG-3

XX PS Example 10; Page 42; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 21) corresponds to amino acid residues  
 CC 72-81 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine

CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX Sequence 10 AA;

Query Match 60.0%; Score 63; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.035;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LNCRCRCGAR 15  
 Db 1 lncrcrcgar 10

RESULT 5

Y06044  
 ID Y06044 standard; Peptide; 9 AA.

XX AC Y06044;

XX DT 16-AUG-1999 (first entry)

XX DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX OS Homo sapiens.

XX PN WO9918206-A2.

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX PA (USSR ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenberg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX PT Cancer antigen NY ESO1/CAG-3

XX PS Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 27) corresponds to amino acid residues  
 CC 73-81 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 9 AA;

SQ

Query Match 56.2%; Score 59; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 NGCCRCGAR 15  
 DB 1 ngccrcgar 9

## RESULT 6

ID Y06060 standard; Peptide; 10 AA.  
 XX Y06060;  
 DT 16-AUG-1999 (first entry)  
 DE Human cancer antigen NY ESO-1/CAG-3 peptide ESO10-68.  
 XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; cytotoxic T lymphocyte; CTL.

OS Homo sapiens.

PN W09918206-A2.

XX 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Rosenberg SA, Wang RF;

DR WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 45; 88pp; English.

CC Peptide ESO10-68 corresponds to amino acid residues 68-77 of  
 CC human NY ESO-1/CAG-3 ORF1 (see Y05965), a new and potent tumour  
 CC antigen capable of eliciting an antigen specific immune response  
 CC by T cells. It was examined for reactivity to a cytotoxic T  
 CC lymphocyte (CTL), measured as release of granulocyte macrophage  
 CC colony stimulating factor. Cancer peptides (see Y05967-87) derived  
 CC from CAG-3, portions of CAG-3 and their variants, are useful as  
 CC cancer vaccines. A claimed method of preventing or inhibiting  
 CC cancer involves administering a cancer peptide, with or without an  
 CC HLA molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 10 AA;

Query Match 54.3%; Score 57; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AASGLNGCCR 11  
 DB 1 aasgngccr 10

Db 1 aasgngccr 10

## RESULT 7

ID Y05994 standard; Peptide; 10 AA.

XX Y05994;

XX 16-AUG-1999 (first entry)

XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

OS Homo sapiens.

PN W09918206-A2.

XX 15-APR-1999.

XX 21-SEP-1998; 98WO-US19609.

XX 08-OCT-1997; 97US-0061428.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Rosenberg SA, Wang RF;

DR WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 42; 88pp; English.

CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 7) corresponds to amino acid residues  
 CC 68-77 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 10 AA;

Query Match 54.3%; Score 57; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AASGLNGCCR 11

DB 1 aasgngccr 10

RESULT 8  
 Y06047  
 ID Y06047 standard; Peptide; 9 AA.

```

XX AC Y06047;
XX 16-AUG-1999 (first entry)
XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
KW vaccine; human leukocyte antigen; HLA.
XX Homo sapiens.
OS
XX W09918206-A2.
PN
XX 15-APR-1999.
PD
XX 21-SEP-1998; 98WO-US19609.
PF
XX 08-OCT-1997; 97US-0061428.
PR
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Rosenberg SA, Wang RF;
PI
XX WPI; 1999-277270/23.
PT
XX Cancer antigen NY ESO1/CAG-3
XX Example 10; Page 43; 88pp; English.
XX This peptide was identified as an HLA peptide motif following a
CC screen for epitopes from the coding region of human ESO-1/CAG-3
CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.
CC The present peptide (ranked 30) corresponds to amino acid residues
CC 69-77 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent
CC tumour antigen capable of eliciting an antigen specific immune
CC response by T cells. Cancer peptides (see Y05967-87) derived from
CC CAG-3, portions of CAG-3 and their variants, are useful as cancer
CC vaccines. A claimed method of preventing or inhibiting cancer
CC involves administering a cancer peptide, with or without an HLA
CC molecule. The cancer peptides form part of, or are derived
CC from, cancers such as primary or metastatic melanoma, thymoma,
CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine
CC cancer, cervical cancer, bladder cancer, kidney cancer and
CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and
CC thyroid cancers.
XX Sequence 9 AA;
SQ
Query Match 50.5%; Score 53; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 ASGLNGCCR 11
DB 1 asgIngccr 9
RESULT 9
Y06059
ID Y06059 standard; Peptide; 10 AA.
XX
XX Y06059;
XX
XX 16-AUG-1999 (first entry)
XX Human cancer antigen NY ESO-1/CAG-3 peptide ESO10-77.
DE

```

```

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
KW vaccine; cytotoxic T lymphocyte; CTL.
XX Homo sapiens.
OS
XX W09918206-A2.
PN
XX 15-APR-1999.
PD
XX 21-SEP-1998; 98WO-US19609.
PF
XX 08-OCT-1997; 97US-0061428.
PR
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Rosenberg SA, Wang RF;
PI
XX WPI; 1999-277270/23.
PT
XX Cancer antigen NY ESO1/CAG-3
XX Example 10; Page 45; 88pp; English.
XX Peptide ESO10-77 corresponds to amino acid residues 77-86 of
CC human NY ESO-1/CAG-3 ORF1 (see Y05965), a new and potent tumour
CC antigen capable of eliciting an antigen specific immune response
CC by T cells. It was examined for reactivity to a cytotoxic T
CC lymphocyte (CTL), measured as release of granulocyte macrophage
CC colony stimulating factor. Cancer peptides (see Y05967-87) derived
CC from CAG-3, portions of CAG-3 and their variants, are useful as
CC cancer vaccines. A claimed method of preventing or inhibiting
CC cancer involves administering a cancer peptide, with or without an
CC HLA molecule. The cancer peptides form part of, or are derived
CC from, cancers such as primary or metastatic melanoma, thymoma,
CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine
CC cancer, cervical cancer, bladder cancer, kidney cancer and
CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and
CC thyroid cancers.
XX Sequence 10 AA;
SQ
Query Match 44.8%; Score 47; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 RCGARGPE 18
DB 1 rcgargpe 8
RESULT 10
Y05993
ID Y05993 standard; Peptide; 10 AA.
XX
XX Y05993;
XX
XX 16-AUG-1999 (first entry)
XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.
DE
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;
KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;
KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;
KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;
KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;
KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;
KW

```

KW vaccine; human leukocyte antigen; HLA.  
 XX Homo sapiens.  
 OS  
 PN W09918206-A2.  
 XX  
 XX 15-APR-1999.  
 XX  
 XX 21-SEP-1998; 98WO-US19609.  
 PF  
 XX 08-OCT-1997; 97US-0061428.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Rosenberg SA, Wang RF;  
 XX WPI; 1999-277270/23.  
 XX  
 XX Cancer antigen NY ESO1/CAG-3  
 PT  
 XX Example 10; Page 42; 88pp; English.  
 PS  
 XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 6) corresponds to amino acid residues  
 CC 77-86 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 XX Sequence 10 AA;  
 SQ

Query Match 44.8%; Score 47; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RCGARGPE 18  
 DB 1 rcgargpe 8  
 |||||

RESULT 11  
 R69313  
 ID R69313 standard; peptide; 10 AA.  
 XX  
 AC R69313;  
 XX  
 XX 24-JUN-1995 (first entry)  
 DT  
 XX Gp IIb/IIIa receptor ligand used in scintigraphic imaging of thrombi.  
 DE  
 XX Scintigraphy; thrombus; thrombi; imaging; specific binding;  
 KW technetium-99m; radiolabelled; Gp IIb/IIIa receptor ligand.  
 XX  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FT Modified-site 2 /note= "S-acetamidomethyl-Cys"  
 FT Modified-site 4  
 FT Modified-site 10 /note= "S-acetamidomethyl-Cys"  
 FT Modified-site 10 /note= "this residue is joined via the OH group

FT to tris(2-succinimidoethyl)amine (TSEA)  
 FT to form one of three such peptides  
 FT connected to the amine; i.e. forming a  
 XX product of formula (peptide)3-TSEA"  
 XX  
 XX W09323085-A.  
 XX  
 XX 25-NOV-1993.  
 PD  
 XX 21-MAY-1993; 93WO-US04794.  
 PF  
 XX 21-MAY-1992; 92US-0886752.  
 PR  
 XX (DIAT-) DIATECH INC.  
 PA  
 XX Dean RT, Lister-James J;  
 XX WPI; 1993-386229/48.  
 DR  
 XX  
 XX Reagent for scintigraphic imaging of thrombi with 99m technetium  
 PT - comprises synthetic peptide which binds to thrombus, covalently  
 PT coupled to metal binding gp.; rapidly cleared from blood and  
 PT tissue  
 XX  
 XX Claim 19; Page 42; 61pp; English.  
 PS  
 XX The invention relates to reagents for scintigraphic imaging of a  
 CC thrombus in-vivo, comprising (A) a specific binding compound capable of  
 CC binding to at least one component of a thrombus, covalently linked  
 CC to (B) a technetium-99m-binding moiety. Specific peptides  
 CC constituting the reagents are claimed as new. The present peptide is  
 CC one such peptide, in which the Cys(Acm)-gly-Cys(Acm) moiety is the  
 CC 99m-Tc binding moiety and the residue constitutes the thrombus-  
 CC binding component.  
 XX  
 XX Sequence 10 AA;  
 SQ

Query Match 40.0%; Score 42; DB 14; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 CCRGARG 16  
 DB 1 ccgcggrg 8  
 |||||

RESULT 12  
 W52088  
 ID W52088 standard; peptide; 10 AA.  
 XX  
 AC W52088;  
 XX  
 XX 23-SEP-1998 (first entry)  
 DT  
 XX Targetting peptide #16 useful as component of thrombolytic agent.  
 DE  
 XX Thrombolytic agent; thrombolytic proteinase; blood clot; fibrin;  
 KW thrombus; antithrombotic activity.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FT Modified-site 2 /note= "The side-chain thiol of Cys is protected by  
 FT Modified-site 4 an acetamidomethyl group"  
 FT Modified-site 4 /note= "The side-chain thiol group of Cys is protected  
 FT by an acetamidomethyl group"  
 XX  
 XX W09824917-A1.  
 PN  
 XX 11-JUN-1998.  
 PD

XX 02-DEC-1997; 97WO-US21918.  
 PF  
 XX 02-DEC-1996; 96US-0753781.  
 PR  
 XX (DIAT-) DIATIDE INC.  
 PA  
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 PA  
 XX Bush LR, Flores-Sanchez S E, Markland FS, Swenson S;  
 PI  
 XX WPI; 1998-333336/29.  
 DR  
 XX New thrombolytic agents - comprise thrombolytic proteinase  
 XX covalently linked to targetting compound for binding to component of  
 PT thrombus  
 PT  
 XX  
 PS Claim 10; Page 64; 79pp; English.  
 XX  
 CC The invention relates to new thrombolytic agents which comprise a  
 CC thrombolytic proteinase covalently linked to a targetting compound  
 CC capable of specifically binding to a component of a thrombus. The  
 CC thrombolytic agents can be used for eliminating thrombi in vivo in, e.g.  
 CC myocardial infarction, cerebral ischaemia, deep vein thrombosis or  
 CC pulmonary embolism. A labelled form of the thrombolytic agent can also  
 CC be used to image thrombi for diagnostic purposes. The thrombolytic  
 CC agents are specifically targetted to thrombus sites in vivo and have  
 CC minimal haemorrhagic side effects and side effects related to non-  
 CC specific proteolysis. The present sequence represents a specifically  
 CC claimed targetting peptide.  
 CC  
 XX Sequence 10 AA;  
 SQ

Query Match 40.0%; Score 42; DB 19; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 CCRGARG 16  
 II II II  
 Db 1 ccgcggrg 8

RESULT 13  
 Y95511  
 ID Y95511 standard; Peptide; 10 AA.  
 XX  
 AC Y95511;  
 XX  
 DT 10-OCT-2000 (first entry)  
 XX  
 DE GPIIb/IIIa receptor ligand with polyvalent linking moiety.  
 XX  
 KW Thrombus; imaging; radioimaging; thrombosis; thromboembolism;  
 KW embolism; diagnosis; technetium-99m; GPIIb/IIIa receptor ligand.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 2  
 FT Modified-site 4 /note= "Cys(ACM)"  
 FT Modified-site 4 /note= "Cys(ACM)"  
 FT Peptide 1..10  
 FT /note= "3 peptides of this sequence are joined  
 FT C-terminally to the polyvalent linking  
 FT moiety tris(succinimidylethyl)amine"  
 XX  
 XX USG083481-A.  
 PN  
 XX 04-JUL-2000.  
 PD  
 XX 27-AUG-1998; 98US-0141127.  
 PF  
 XX

PR 21-MAY-1992; 92US-0886752.  
 PR 21-MAY-1993; 93WO-US04794.  
 PR 05-JAN-1995; 95US-0335832.  
 XX  
 PA (DIAT-) DIATIDE INC.  
 XX  
 PI Lister-James J, Dean RT;  
 XX  
 DR WPI; 2000-498061/44.  
 XX  
 CC Composition comprises technetium-99m and polyamide reagent which binds  
 CC to thrombi, useful as scintigraphic imaging agent for imaging sites of  
 PT thrombus formation in vivo  
 PT  
 XX  
 PS Disclosure; Column 10; 27pp; English.  
 XX  
 CC The present sequence is that of a GPIIb/IIIa receptor ligand. The  
 CC invention relates to radiolabeled reagents that are scintigraphic  
 CC imaging agents for imaging sites of thrombus formation in vivo.  
 CC The reagents each comprise a specific binding compound, such as the  
 CC present peptide, that is capable of binding to at least 1 component  
 CC of a thrombus, and which is covalently linked to a radiolabel-binding  
 CC moiety and may include a polyvalent linking moiety. A method for  
 CC using such a reagent labeled with technetium-99m to image a thrombus  
 CC site in a mammalian body is claimed.  
 XX  
 SQ Sequence 10 AA;  
 -

Query Match 40.0%; Score 42; DB 21; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 CCRGARG 16  
 II II II  
 Db 1 ccgcggrg 8

RESULT 14  
 Y97091  
 ID Y97091 standard; peptide; 11 AA.  
 XX  
 AC Y97091;  
 XX  
 DT 22-DEC-2000 (first entry)  
 XX  
 DE Tau conotoxin sequence.  
 XX  
 KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;  
 KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;  
 KW migraine; treatment; therapy.  
 XX  
 OS Conus marmoreus.  
 XX  
 PN WO200046371-A1.  
 XX  
 PD 10-AUG-2000.  
 XX  
 PF 04-FEB-2000; 2000WO-US03021.  
 XX  
 PR 04-FEB-1999; 99US-0118642.  
 XX  
 PA (UTAH ) UNIV UTAH RES FOUND.  
 PA (COGN-) COGNETIX INC.  
 XX  
 PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;  
 PI Jones RM;  
 XX  
 DR WPI; 2000-543489/49.  
 XX  
 PT Novel analgesic short peptides from predatory cone snails for treating  
 PT acute, chronic and neuropathic pain and migraines  
 XX

PS Claim 2; Page 28; 58pp; English.

XX The venom of predatory cone snails comprises relatively small  
CC peptides which are targetted to various neuromuscular receptors  
CC and may be equivalent in their pharmacological diversity to the  
CC alkaloids of plants or secondary metabolites of microorganisms.  
CC Several peptides have been characterised from Conus venoms. These  
CC include the alpha and mu conotoxins which target nicotinic  
CC acetylcholine receptors and muscle sodium channels respectively.  
CC Chronic or intractable pain as well as neuropathic pain is  
CC currently treated with a range of analgesic compounds. It is  
CC thought that the tau-conotoxins described could have applications  
CC as analgesic drugs and could be used for treating such acute,  
CC chronic and neuropathic pain as well as migraines.

XX Sequence 11 AA;

Query Match 39.0%; Score 41; DB 21; Length 11;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 NGCCRCG 13  
| | | | |  
Db 1 ngccrag 7

RESULT 15

Y971109  
ID Y97109 standard; peptide; 11 AA.

XX AC Y97109;

XX DT 22-DEC-2000 (first entry)

XX DE Tau conotoxin sequence.

XX Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;  
KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;  
KW migraine; treatment; therapy.

XX OS Conus marmoreus.

XX PN WO200046371-A1.

XX PD 10-AUG-2000.

XX PF 04-FEB-2000; 2000WO-US03021.

XX PR 04-FEB-1999; 99US-0118642.

XX PA (UTAH ) UNIV UTAH RES FOUND.

XX PA (COGN-) COGNETIX INC.

XX PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;  
PI Jones RM;

XX DR WPI; 2000-543489/49.

XX PT Novel analgesic short peptides from predatory cone snails for treating  
PT acute, chronic and neuropathic pain and migraines

XX PS Claim 25; Page 29; 58pp; English.

XX The venom of predatory cone snails comprises relatively small  
CC peptides which are targetted to various neuromuscular receptors  
CC and may be equivalent in their pharmacological diversity to the  
CC alkaloids of plants or secondary metabolites of microorganisms.  
CC Several peptides have been characterised from Conus venoms. These  
CC include the alpha and mu conotoxins which target nicotinic  
CC acetylcholine receptors and muscle sodium channels respectively.  
CC Chronic or intractable pain as well as neuropathic pain is  
CC currently treated with a range of analgesic compounds. It is

CC thought that the tau-conotoxins described could have applications  
CC as analgesic drugs and could be used for treating such acute,  
CC chronic and neuropathic pain as well as migraines.

XX Sequence 11 AA;

Query Match 39.0%; Score 41; DB 21; Length 11;  
Best Local Similarity 85.7%; Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 NGCCRCG 13  
| | | | |  
Db 1 ngccrag 7

Search completed: May 8, 2001, 15:07:10  
Job time: 226 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:08 ; Search time 40.07 Seconds  
(without alignments)  
15.388 Million cell updates/sec

Title: US-09-165-546A-11

Perfect score: 105

Sequence: 1 GAASGLNGCCRCGARGPE 18

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34.5	32.9	16	CXAL_CONEP	P56638 conus episc
2	33.5	31.9	16	CXAA_CONPE	P50984 conus penna
3	31	29.5	25	LYC_ASTRU	P37715 asterias ru
4	30	28.6	19	CX2_CONST	P28879 conus stria
5	29.5	28.1	15	CX2_CONAL	P56640 conus aulic
6	28.5	27.1	12	CX1_CONIM	P50983 conus imper
7	28	26.7	25	MT_AGABI	P04358 agaricus bi
8	27.5	26.2	16	CXAB_CONPE	P50985 conus penna
9	27.5	26.2	25	MT_NEUCR	P02807 neuropsora
10	26	24.8	19	LCRP_PETMA	Q10996 petromyzon
11	25	23.8	15	TAL_TREBR	P34070 tremella br
12	25	23.8	17	FLAW_AZOC	P23002 azotobacter
13	25	23.8	20	FLAW_AZOV	P52964 azotobacter
14	24.5	23.3	22	ANFC_CHICK	P21805 gallus gall
15	24.5	23.3	24	CXOA_CONST	P28880 conus stria
16	24	22.9	16	CX1_CONAL	P56639 conus aulic
17	24	22.9	16	CX2_CONMA	P56636 conus magus
18	24	22.9	16	CX3_CONAL	P56641 conus aulic
19	24	22.9	20	PSAE_PEA	P20119 pisum sativ
20	24	22.9	23	CHIB_PEA	P21227 pisum sativ
21	24	22.9	24	PSAC_CUGSA	P42046 cucumis sat
22	24	22.9	25	CX4_CONPU	P55963 conus purpu
23	24	22.9	25	CXOA_CONMA	P05484 conus magus
24	24	22.9	25	G3P2_JACOR	P80447 jaculus ori
25	24	22.9	25	K6P2_THETH	P21778 thermus aqu
26	24	22.9	25	LEO1_BIOGL	P80742 biomphalari
27	23	21.9	13	CX1_CONST	P15471 conus stria
28	23	21.9	15	DIDH_PSESP	P80701 pseudomonas
29	23	21.9	22	LANM_STRMU	P80666 streptococc
30	22	21.0	13	CX2_CONGE	P01520 conus geogr
31	22	21.0	13	CXAA_CONST	P28878 conus stria
32	22	21.0	14	CX1_CONCN	P56973 conus conso
33	22	21.0	14	CXAL_CONMA	P01521 conus magus

34	22	21.0	15	1	CXAL_CONGE	P01519 conus geogr
35	22	21.0	16	1	BAL_EUBSP	P32371 eubacterium
36	22	21.0	19	1	MDH_SHEPU	P82177 shewanella
37	22	21.0	21	1	FER_PYRWO	P81638 pyrococcus
38	22	21.0	24	1	VORC_METTM	P80909 methanobact
39	21.5	20.5	18	1	CX1_CONER	P50982 conus ermin
40	21	20.0	12	1	RR16_GINBI	P36207 ginkgo bilbo
41	21	20.0	13	1	CXET_CONTE	P81755 conus texti
42	21	20.0	15	1	UC06_MAIZE	P80612 zea mays (m
43	21	20.0	23	1	FMK7_PSEAE	Q53391 pseudomonas
44	21	20.0	25	1	RR16_SPIOL	P28807 spinacia ol
45	20	19.0	19	1	HMD_METWO	P32441 methanobact

#### ALIGNMENTS

##### RESULT 1

CXAL\_CONEP STANDARD; PRT; 16 AA.  
AC P56638;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE ALPHA-CONOTOXIN EPI.  
OS Conus episcopatius (Bishop's cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=88764;  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).  
RX MEDLINE=98376423; PubMed=9708977;  
RA Hu S.H., Loughnan M., Miller R., Weeks C.M., Blessing R.H.,  
RA Alewood P.F., Lewis R.J., Martin J.L.;  
RT "The 1.1-A resolution crystal structure of [Tyr15]Epi, a novel  
alpha-conotoxin from Conus episcopatius, solved by direct methods.";  
RL Biochemistry 37:11425-11433(1998).  
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC  
CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-2 AND ALPHA-  
CC 3/BETA-4 SUBUNITS.  
CC PDB: 1A0M; 13-JAN-99.  
DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;  
KW Venom; Sulfatation; 3D-structure.  
FT DISULFID 2 8  
FT DISULFID 3 16  
FT MOD\_RES 15 15 SULFATATION.  
FT MOD\_RES 16 16 AMIDATION.  
SQ SEQUENCE 16 AA; 1792 MW; C63385F376C99B4C CRC64;

Query Match 32.9%; Score 34.5; DB 1; Length 16;  
Best Local Similarity 42.9%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

QY 8 GCC---RCGARGPE 18  
||| || |  
Db 1 GCCSDPRCWNMPD 14

##### RESULT 2

CXAA\_CONPE STANDARD; PRT; 16 AA.  
AC P50984;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE ALPHA-CONOTOXIN PNIA.  
OS Conus pennaceus.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=37335;

```
CC -1- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.
DR PIR; A11762; A11762.
DR InterPro; IPR001916; -.
DR PROSITE; PS00128: LACTALBUMIN_LYSOZYME: PARTIAL.
DR Hydrolase; Glycosidase; Bacteriolytic enzyme.
FW NON_TER      25
SQ SEQUENCE     25 AA; 2568 MW; 6FA1D6BE87C5BE18 CRC64;

Query Match          29.5%; Score 31; DB 1; Length 25;
Best Local Similarity 41.2%; Pred.No. 2e+02;
Matches       7; Conservative   1; Mismatches    3; Indels    6; Gaps    1;

QY      7 NGCCRC-----GARGP 17
        :|||:|||         |||
DB      6 SGCLRLCIVVSGXRMP 22

RESULT 4
CXA2_CONST
ID   CXA2_CONST          STANDARD;              PRT;           19 AA.
AC   P28879;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   ALPHA-CONOTOXIN SII (S2).
OS   Conus striatus (Striated cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
NCBI_TaxID=6493;
RN   [1]
RS   SEQUENCE, AND SYNTHESIS.
RC   TISSUE-Venom:
RX   MEDLINE=93003172; PubMed=1390774;
RA   Ramilo C., Zafaralla G.C., Nadasdi L., Hammerland L.G., Yoshikami D.,
RA   Gray W.R., Kristipati R., Ramachandran J., Miljanich G., Olivera B.M.,
RA   Cruz L.J.;
RA   "Novel alpha- and omega-conotoxins from Conus striatus venom.";
RL   Biochemistry 31:9919-9926(1992).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC CC
PIR; A44379; A44379.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Venom.
FT DISULFID      2      18      POTENTIAL.
FT DISULFID      3      18      BY SIMILARITY.
FT DISULFID      4      14      BY SIMILARITY.
SQ SEQUENCE     19 AA; 1797 MW; A15B024273EF4CF CRC64;

Query Match          28.6%; Score 30; DB 1; Length 19;
Best Local Similarity 60.0%; Pred.No. 2.2e+02;
Matches       6; Conservative   0; Mismatches    4; Indels    0; Gaps    0;

QY      8 GCCRCGARGP 17
        |||||         |||
DB      1 GCCCNPCGP 10

RESULT 5
CXA2_CONAL
ID   CXA2_CONAL          STANDARD;              PRT;           15 AA.
AC   P56640;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ALPHA-CONOTOXIN AUTOB.
OS   Conus aulicus (Court cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
NCBI_TaxID=89437;
RN   [1]
RS   SEQUENCE, AND SYNTHESIS.
RC
RX
RA
RA
```

```
RC TISSUE-Venom;
RX MEDLINE-99003392; PubMed-9786965;
RA Luo S., Kulak J.M., Cartier G.E., Jacobsen R.B., Yoshikami D.,
RA Olivera B.M., McIntosh J.M.;
RT "Alpha-conotoxin AuIB selectively blocks alpha3 beta4 nicotinic
RL acetylcholine receptors and nicotine-evoked norepinephrine release.";
RL J. Neurosci. 18:8571-8579(1998).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. THIS PEPTIDE BLOCKS MAMMALIAN NICOTINIC
CC ACETYLCHOLINE RECEPTORS COMPOSED OF ALPHA-3/BETA-4 SUBUNITS.
CC -1- MASS SPECTROMETRY: MW-1572.5; METHOD-ELECTROSPRAY.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 2 8
FT DISULFID 3 15
FT MOD_RES 15 15
FT MOD_RES 15 15
SQ SEQUENCE 15 AA; 1578 MW; 84EFE95FDC700155 CRC64;

Query Match 28.1%; Score 29.5; DB 1; Length 15;
Best Local Similarity 42.9%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

QY 8 GCCR---CGARGPE 18
D 111 111
Db 1 GCCSYPPCFATNPD 14

RESULT 6
CXAL_CONTIM STANDARD; PRT; 12 AA.
AC P50983;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ALPHA-CONOTOXIN IMI.
OS Conus imperialis (Imperial cone).
OS Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=35631;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE-Venom;
RX MEDLINE-94266889; PubMed-8206995;
RA McIntosh J.M., Yoshikami D., Mahe E., Nielsen D.B., Rivier J.E.,
RA Gray W.R., Olivera B.M.;
RT "A nicotinic acetylcholine receptor ligand of unique specificity,
RT alpha-conotoxin Imi.";
RL J. Biol. Chem. 269:16733-16739(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE-95379776; PubMed-7651351;
RA Johnson D.S., Martinez J., Elgoyhen A.B., Heinemann S.F.,
RA McIntosh J.M.;
RT "Alpha-conotoxin Imi exhibits subtype-specific nicotinic
RT acetylcholine receptor blockade: preferential inhibition of homomeric
RT alpha 7 and alpha 9 receptors.";
RL Mol. Pharmacol. 48:194-199(1995).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE-99212205; PubMed-10194298;
RA Rogers J.P., Luginbuhl P., Shen G.S., McCabe R.T., Stevens R.C.,
RA Wemmer D.E.;
RT "NMR solution structure of alpha-conotoxin ImI and comparison to
RT other conotoxins specific for neuronal nicotinic acetylcholine
RT receptors.";
RL Biochemistry 38:3874-3882(1999).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE-99158061; PubMed-10050774;
RA Meshinnikov I.V., Shenkarev Z.O., Zhmak M.N., Ivanov V.T.,
RA Methfessel C., Tsetlin V.I., Arseniev A.S.;
```

```
*NMR spatial structure of alpha-conotoxin ImI reveals a common
scaffold in snail and snake toxins recognizing neuronal nicotinic
acetylcholine receptors.*;
FEBS Lett. 444:275-280(1999).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE-99324017; PubMed-10395477;
RA Gehrmann J., Daly N.B., Alewood P.F., Craik D.J.;
RT "Solution structure of alpha-conotoxin ImI by 1H nuclear magnetic
RT resonance.*;
RL J. Med. Chem. 42:2364-2372(1999).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. IT IS HIGHLY ACTIVE AGAINST THE NEUROMUSCULAR
CC RECEPTOR IN FROG BUT NOT IN MICE. IN CONTRAST, IT INDUCES SEIZURES
CC WHEN INJECTED CENTRALLY IN MICE AND RATS. IT TARGETS NEURONAL
CC NACHRS IN MAMMALS. BLOCKS HOMOMERIC ALPHA-7 NICOTINIC RECEPTORS
CC WITH THE HIGHEST APPARENT AFFINITY AND HOMOMERIC ALPHA-9 RECEPTORS
CC WITH 8-FOLD LOWER AFFINITY. HAS NO EFFECT ON RECEPTORS COMPOSED OF
CC ALPHA-2/BETA-2, ALPHA-3/BETA-4, OR ALPHA-4/BETA-2, ALPHA-2/BETA-4,
CC ALPHA-3/BETA-4, OR ALPHA-4/BETA-4 SUBUNIT COMBINATIONS.
DR PDB: 1IM1; 15-JUN-99.
DR PDB: 1IM1; 23-APR-99.
DR PDB: 1CNL; 27-MAY-99.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 12
FT MOD_RES 12 12
SQ SEQUENCE 12 AA; 1357 MW; 9C29CEA545A4176A CRC64;

Query Match 27.1%; Score 28.5; DB 1; Length 12;
Best Local Similarity 54.5%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 8 GCC---RCGAR 15
D 111 111
Db 1 GCCSDPRCAWR 11

RESULT 7
MT_AGABI STANDARD; PRT; 25 AA.
AC P04358;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN (MT).
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;
OC Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN [1]
RP SEQUENCE.
RC STRAIN=A-32; TISSUE=Mycelium;
RA Munger K., Lerch K.;
RT "Copper metallothionein from the fungus Agaricus bisporus: chemical
RT and spectroscopic properties.";
RL Biochemistry 24:6751-6756(1985).
CC -1- MISCELLANEOUS: THE SEVEN CYSTEINES BIND SIX COPPER (CUPROUS) IONS.
CC -1- MISCELLANEOUS: THE ABSORPTION, LUMINESCENT, AND STEREOTYPICAL
CC PROPERTIES OF THE COPPER MT ARE ATTRIBUTED TO THE METAL-THIOLATE
CC COMPLEX BECAUSE THEY ARE NOT PRESENT IN THE APOPROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 8 IN METALLOTHIONEIN SUPERFAMILY.
DR PIR: A03286; SMR.
KW Metal-binding; Metal-thiolate cluster; Chelation; Copper.
SQ SEQUENCE 25 AA; 2233 MW; 33AB70F21023CFF7 CRC64;

Query Match 26.7%; Score 28; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```

QY 3 ASLNGCCRCG 13
Db 14 ASGQCTSCG 24

RESULT 8
CXAB_CONPE
ID CXAB_CONPE STANDARD; PRT; 16 AA.
AC P50985;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE ALPHA-CONOTOXIN PNIB.
OS Conus pennaceus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=37335;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=94347719; PubMed=8068627;
RA Fainzilber M., Hasson A., Oren R., Burlingame A.L., Gordon D.,
RA Spira M.E., Zlotkin E.;
RT "New-mollusc-specific alpha-conotoxins block Aplysia neuronal
RT acetylcholine receptors.";
RL Biochemistry 33:9523-9529(1994).
RN [2]
RP SULFATATION OF TYR-15.
RX MEDLINE=99242956; PubMed=10226369;
RA Wolfender J.L., Chu F., Ball H., Wolfender F., Fainzilber M.,
RA Baldwin M.A., Burlingame A.L.;
RT "Identification of tyrosine sulfation in Conus pennaceus conotoxins
RT alpha-PnIA and alpha-PnIB: further investigation of labile sulfo- and
RT phosphopeptides by electrospray, matrix-assisted laser
RT desorption/ionization (MALDI), and atmospheric pressure MALDI mass
RT spectrometry.";
RL J. Mass Spectrom. 34:447-454(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
RX MEDLINE=97444322; PubMed=9298951;
RA Hu S.H., Gehrmann J., Alewood P.F., Craik D.J., Martin J.L.;
RT "Crystal structure at 1.1-A resolution of alpha-conotoxin PnIB:
RT comparison with alpha-conotoxins PnIA and GI.";
RL Biochemistry 36:11323-11330(1997).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. IN CONTRAST TO OTHER ALPHA-CONOTOXINS, WHICH ARE
CC SELECTIVE FOR VERTEBRATE SKELETAL MUSCLE NACHR, THE CONUS
CC PENNACEUS ALPHA-CONOTOXINS BLOCK NACHR IN MOLLUSCS.
DR PDB; 1AKG; 20-MAY-98.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Sulfatation; Venom; 3D-structure.
FT DISULFID 2 8
FT DISULFID 3 16
FT MOD_RES 15 15 SULFATATION.
FT MOD_RES 16 16 AMIDATION.
SQ SEQUENCE 16 AA; 1643 MW; 05310FF95ED86AF5 CRC64;

Query Match 26.2%; Score 27.5; DB 1; Length 16;
Best Local Similarity 35.7%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

QY 8 GCCR---CGARGE 18
Db 1 GCCSLPPCALSNPD 14

RESULT 9
MT_NEUCR
ID MT_NEUCR STANDARD; PRT; 25 AA.
AC P02807;

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-JUL-1999 (Rel. 38, Last annotation update)
DE METALLOTHIONEIN (MT).
GN CMT.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86030247; PubMed=2932331;
RA Mungler K., Germann U.A., Lerch K.;
RT "Isolation and structural organization of the Neurospora crassa
RT copper metallothionein gene.";
RL EMBO J. 4:2665-2668(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91112714; PubMed=2148862;
RA Romeyer F.M., Jacobs F.A., Brousseau R.;
RT "Expression of a Neurospora crassa metallothionein and its variants
RT in Escherichia coli.";
RL Appl. Environ. Microbiol. 56:2748-2754(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88029908; PubMed=2959528;
RA Muenger K., Germann U.A., Lerch K.;
RT "Isolation and regulation of expression of the Neurospora crassa
RT copper metallothionein gene.";
RL Experientia Suppl. 52:393-400(1987).
RN [4]
RP SEQUENCE.
RX MEDLINE=80143244; PubMed=6444697;
RA Lerch K.;
RT "Copper metallothionein, a copper-binding protein from Neurospora
RT crassa.";
RL Nature 284:368-370(1980).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE=89302887; PubMed=2525920;
RA Malikayil J.A., Lerch K., Armitage I.M.;
RT "Proton NMR studies of a metallothionein from Neurospora crassa:
RT sequence-specific assignments by NOE measurements in the rotating
RT frame.";
RL Biochemistry 28:2991-2995(1989).
CC -1- INDUCTION: BY COPPER.
CC -1- MISCELLANEOUS: THE SEVEN CYSTEINES BIND SIX COPPER (CUPROUS) IONS.
CC -1- SIMILARITY: BELONGS TO FAMILY 8 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03009; CAA26793.1; -
DR EMBL; M59836; AAA33594.1; -
DR EMBL; M27709; AAA33595.1; -
DR PIR; A24641; SMNC.
KW Metal-binding; Metal-thiolate cluster; Chelation; Copper.
FT INIT_MET 0 0
FT METAL 3 3 COPPER.
FT METAL 5 5 COPPER.
FT METAL 11 11 COPPER.
FT METAL 13 13 COPPER.
FT METAL 17 17 COPPER.
FT METAL 19 19 COPPER.
FT METAL 22 22 COPPER.
FT CONFLICT 21 21 N->T (IN REF. 2).
SQ SEQUENCE 25 AA; 2234 MW; 256CB6335FD422FC CRC64;

```

Query Match 26.2%; Score 27.5; DB 1; Length 25;  
 Best Local Similarity 41.7%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

OY 8 GC-----CRCGA 14  
 II I I I  
 DB 4 GCSGASSNCGS 15

RESULT 10  
 LCRP\_PETMA STANDARD; PRT; 19 AA.  
 AC Q10996;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CORTICOSTATIN-RELATED PROTEIN LCRP.  
 OS Petromyzon marinus (Sea lamprey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
 OX NCBI\_TaxID=7757;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin;  
 RX MEDLINE=96321324; PubMed=8759287;  
 RA Conlon J.M., Sower S.A.;  
 RT "Isolation of a peptide structurally related to mammalian  
 RT corticostatins from the lamprey Petromyzon marinus.";  
 RL Comp. Biochem. Physiol. 114B:133-137(1996).  
 CC -!- FUNCTION: MAY HAVE MICROBICIDAL ACTIVITIES. MAY INHIBIT  
 CC ACTIONS OF THE CORTICOSTATINS.  
 CC -!- MASS SPECTROMETRY: MW=2201; MW ERR=0.4; METHOD=ELECTROSPRAY.  
 CC -!- SIMILARITY: BELONGS TO THE FAMILY OF CORTICOSTATIN/DEFENSIN  
 CC PEPTIDES.  
 KW Defensein; Antibiotic.  
 FT DISULFID 1 18 BY SIMILARITY.  
 FT DISULFID 3 9 BY SIMILARITY.  
 FT DISULFID 8 17 BY SIMILARITY.  
 SQ SEQUENCE 19 AA; 2209 MW; 8D9CEDC71A199AE5 CRC64;

Query Match 24.8%; Score 26; DB 1; Length 19;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 10 CRCGAR 15  
 I I I I  
 DB 1 CPCGRR 6

RESULT 11  
 TAL\_TREBR STANDARD; PRT; 15 AA.  
 AC P34070;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE TREMEROGEN A-I.  
 OS Tremella brasiliensis (Jelly fungus).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Tremellales;  
 OC Tremellaceae; Tremella.  
 OX NCBI\_TaxID=29896;  
 RN [1]  
 RP SEQUENCE.  
 RA Ishibashi Y., Sakagami Y., Isogai A., Suzuki A.;  
 RT "Structures of Tremeregens A-9291-I and A-9291-VIII: peptidyl sex  
 RT hormones of Tremella brasiliensis.";  
 RL Biochemistry 23:1399-1404(1984).  
 CC -!- FUNCTION: TREMEROGEN A-I IS PRODUCED BY THE A MATING-TYPE CELLS  
 CC AND INDUCES FORMATION OF CONJUGATION TUBES IN A MATING-TYPE CELLS.  
 KW pheromone; Prenylation; Lipoprotein.

FT LIPID 15 15 FARNESYL.  
 SQ SEQUENCE 15 AA; 1339 MW; 3AABA4FC2D6D05333 CRC64;

Query Match 23.8%; Score 25; DB 1; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 8.9e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

OY 1 GAASGLNGC 9  
 I I I I I  
 DB 9 GASSG--GC 15

RESULT 12  
 FLAW\_AZOCH STANDARD; PRT; 17 AA.  
 ID P23002;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE FLAVODOXIN A (FLDA) (FRAGMENT).  
 OS Azotobacter chroococcum mcd 1.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Azotobacter.  
 OX NCBI\_TaxID=355;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=MCD 1155;  
 RX MEDLINE=91315397; PubMed=1859358;  
 RA Bagby S., Barker P.D., Hill H.A.O., Sanghera G.S., Dunbar B.,  
 RA Ashby G.A., Eady R.R., Thorneley R.N.F.;  
 RT "Direct electrochemistry of two genetically distinct flavodoxins  
 RT isolated from Azotobacter chroococcum grown under nitrogen-fixing  
 RT conditions.";  
 RL Biochem. J. 277:313-319(1991).  
 CC -!- FUNCTION: LOW-POTENTIAL ELECTRON DONOR TO A NUMBER OF REDOX  
 CC ENZYMES.  
 CC -!- COFACTOR: FMN.  
 CC -!- SIMILARITY: BELONGS TO THE FLAVODOXIN FAMILY.  
 DR PIR: S16929; S16929.  
 DR InterPro: IPR001226;  
 DR PROSITE: PS00201; FLAVODOXIN; PARTIAL.  
 KW Electron transport; Flavoprotein; FMN.  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 1692 MW; DE7B6D4A36B2A4C7 CRC64;

Query Match 23.8%; Score 25; DB 1; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 9.8e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAASGLNG 8  
 I I I I I  
 DB 8 GSSSGVTG 15

RESULT 13  
 FLAW\_AZOVI STANDARD; PRT; 20 AA.  
 ID P52964;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE FLAVODOXIN 1 (FRAGMENT).  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Azotobacter.  
 OX NCBI\_TaxID=354;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=UW 136;  
 RX MEDLINE=96276406; PubMed=8694750;  
 RA Gangeswaran R., Eady R.R.;

```

RT "Flavodoxin 1 of Azotobacter vinelandii: characterization and role in
RT electron donation to purified assimilatory nitrate reductase.";
RL Biochem. J. 317:103-108(1996).
CC -!- FUNCTION: LOW-POTENTIAL ELECTRON DONOR TO A NUMBER OF REDOX
CC ENZYMES.
CC -!- COFACTOR: FMN.
CC -!- SUBUNIT: MONOMER.
CC -!- MASS SPECTROMETRY: MW=19430; MW_ERR=3; METHOD=ELECTROSPRAY.
CC -!- SIMILARITY: BELONGS TO THE FLAVODOXIN FAMILY.
DR InterPro: IPR001226; -.
DR Pfam: PF00258; flavodoxin; 1.
DR PROSITE: PS00201; FLAVODOXIN: PARTIAL.
KW Electron transport; Flavoprotein; FMN.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2043 MW; 95B3601FAD7D4A36 CRC64;

Query Match 23.8%; Score 25; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAASGLNG 8
Db 8 GSSSGVTG 15

RESULT 14
ANFC_CHKCK STANDARD; PRT; 22 AA.
AC P21805;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE C-TYPE NATRIURETIC PEPTIDE (CNP).
GN NPPC.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91113186; PubMed=1989595;
RA Arimura J.J., Minamino N., Kangawa K., Matsuo H.;
RT "Isolation and identification of C-type natriuretic peptide in
RT chicken brain.";
RL Biochem. Biophys. Res. Commun. 174:142-148(1991).
CC -!- FUNCTION: VASORELAXANT ACTIVITY. HAS A GMP-STIMULATING ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED.
DR PIR: JT0581; JT0581.
DR InterPro: IPR000663; -.
DR Pfam: PF00212; ANP; 1.
DR PRINTS: PR00710; NATPEPTIDES.
DR PROSITE: PS00263; NATRIURETIC_PEPTIDE; 1.
KW Vasoactive.
FT DISULFID 6 22
SQ SEQUENCE 22 AA; 2244 MW; FC2A4706D8DAC025 CRC64;

Query Match 23.3%; Score 24.5; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 GAASGLNGC 9
Db 15 GSMSGL-GC 22

RESULT 15
CXOA_CONST STANDARD; PRT; 24 AA.
ID CXOA_CONST

```

```

AC P28880;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE OMEGA-CONOTOXIN SVIA.
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6493;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Venom;
RX MEDLINE=93003172; PubMed=1390774;
RA Ramilo C., Zafaralla G.C., Nadasdi L., Hammerland L.G., Yoshikami D.,
RA Gray W.R., Kristipati R., Ramachandran J., Miljanich G., Olivera B.M.,
RA Cruz L.J.;
RT "Novel alpha- and omega-conotoxins from Conus striatus venom.";
RL Biochemistry 31:9919-9926(1992).
CC -!- FUNCTION: OMEGA-CONOTOXINS ACT AT PRESYNAPTIC MEMBRANES, THEY BIND
CC AND BLOCK THE CALCIUM CHANNELS.
DR PIR: B44379; B44379.
KW Presynaptic neurotoxin; Calcium channel inhibitor; Venom;
KW Amidation; Hydroxylation.
FT DISULFID 1 15
FT DISULFID 8 18
FT DISULFID 14 23
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 24 24 AMIDATION.
SQ SEQUENCE 24 AA; 2485 MW; B29EFC982ABBE644 CRC64;

Query Match 23.3%; Score 24.5; DB 1; Length 24;
Best Local Similarity 38.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 GAASGLNGCCRCG 13
Db 5 GSPGVTIC-CG 16

Search completed: May 8, 2001, 15:14:08
Job time: 529 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:16:06 ; Search time 114.89 Seconds  
(without alignments)  
18.363 Million cell updates/sec

Title: US-09-165-546A-11  
Perfect score: 105  
Sequence: 1 GAASGLNGCCRCGARGPE 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 6627

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- SPTREMBL15:\*
- 1: sp-archaea:\*
  - 2: sp-bacteria:\*
  - 3: sp-fungi:\*
  - 4: sp-human:\*
  - 5: sp-invertebrate:\*
  - 6: sp-mammal:\*
  - 7: sp-mhc:\*
  - 8: sp-organelle:\*
  - 9: sp-phage:\*
  - 10: sp-plant:\*
  - 11: sp-rodent:\*
  - 12: sp-unclassified:\*
  - 13: sp-vertebrate:\*
  - 14: sp-virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36.5	34.8	22	5	Q17072
2	34	32.4	18	11	Q63136
3	30.5	29.0	24	3	O00059
4	29.5	28.1	16	5	P82706
5	29.5	28.1	18	4	Q14042
6	28	26.7	18	13	Q90790
7	27.5	26.2	25	2	Q9WW14
8	27	25.7	24	6	Q28787
9	27	25.7	24	6	Q28430
10	27	25.7	24	6	Q28781
11	26	24.8	24	11	O89021
12	25	23.8	13	4	P82276
13	25	23.8	18	4	Q9UE38
14	25	23.8	19	2	Q9SLW0
15	25	23.8	21	4	O07101
16	25	23.8	21	4	Q9UMU6
17	25	23.8	21	5	Q9U400
18	25	23.8	24	5	P82176
19	24	22.9	9	5	Q9TWD6

20	24	22.9	10	1	O93739	O93739 methanother
21	24	22.9	10	13	P82084	P82084 limnodynast
22	24	22.9	12	13	P82085	P82085 limnodynast
23	24	22.9	15	8	Q95773	Q95773 conolophus
24	24	22.9	16	4	Q15632	Q15632 homo sapien
25	24	22.9	16	6	Q9TRM0	Q9TRM0 bos taurus
26	24	22.9	18	4	Q9Y6D8	Q9Y6D8 homo sapien
27	24	22.9	18	10	Q9S805	Q9S805 arabidopsis
28	24	22.9	20	11	Q9JKV8	Q9JKV8 mus musculus
29	24	22.9	21	11	O88229	O88229 mus musculus
30	24	22.9	22	2	O52435	O52435 pseudomonas
31	24	22.9	22	6	O9TRG6	O9TRG6 sus scrofa
32	24	22.9	22	11	O88222	O88222 mus musculus
33	24	22.9	23	2	Q04274	Q04274 bordetella
34	24	22.9	23	3	O14431	O14431 cryptonectr
35	24	22.9	23	4	O14606	O14606 homo sapien
36	24	22.9	23	5	O9TJW9	O9TJW9 mytilus edu
37	24	22.9	24	2	O9KIL6	O9KIL6 streptomyce
38	24	22.9	25	6	O9NIV5	O9NIV5 equus cabal
39	24	22.9	25	11	O08800	O08800 mus musculus
40	24	22.9	25	14	O11469	O11469 hepatitis c
41	23.5	22.4	22	3	O13443	O13443 cladosporiu
42	23.5	22.4	22	14	O85481	O85481 rous sarcom
43	23	21.9	13	11	Q9QV13	Q9QV13 cavia (guin
44	23	21.9	15	4	O9Y4Z9	O9Y4Z9 homo sapien
45	23	21.9	15	6	O9TRP2	O9TRP2 sus scrofa

ALIGNMENTS

RESULT 1

Q17072	PRELIMINARY;	PRT;	22 AA.
AC	O17072;		
DT	O1-NOV-1996 (TRENBLrel. 01, Created)		
DT	O1-NOV-1996 (TRENBLrel. 01, Last sequence update)		
DE	O1-NOV-1998 (TRENBLrel. 08, Last annotation update)		
DE	SILKMOTH CHORION MRNA (FRAGMENT).		
OS	Antheraea polyphemus (Polyphemus moth).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;		
OC	Bombycoidea; Saturniidae; Saturniinae; Antheraea.		
OX	NCBI_TaxID=7120;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=78237860; PubMed=277327;		
RA	Sim G.K., Efratiadis A., Jones C.W., Kafatos F.C., Koehler M.,		
RA	Kronenberg H.M., Maniatis T., Regier J.C., Roberts B.F., Rosenthal N.;		
RT	"Studies on the structure of genes expressed during development."		
RL	Cold Spring Harb. Symp. Quant. Biol. 42:933-945(1978).		
DR	EMBL; M10845; AAA27784.1;		
KW	Chorion.		
FT	NON_TER	1	
FT	NON_TER	22	
SQ	SEQUENCE	22 AA;	1944 MW; 14AD250B918D680D CRC64;

Query Match 34.8%; Score 36.5; DB 5; Length 22;  
Best Local Similarity 70.0%; Pred. No. 69;  
Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY	5	GLNGC-CRCG	13
Db	13	GLGGCGCGCG	22

RESULT 2

Q63136	PRELIMINARY;	PRT;	18 AA.
ID	Q63136		
AC	Q63136;		
DT	O1-NOV-1996 (TRENBLrel. 01, Created)		

DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DT RAT ALPHA-CRYSTALLIN B CHAIN.  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91093055; PubMed=2176207;  
 RA Iwaki A., Iwaki T., Goldman J.E., Liem R.K.;  
 RT "Multiple mRNAs of rat brain alpha-crystallin B chain result from  
 alternative transcriptional initiation.";  
 RL J. Biol. Chem. 265:22197-22203(1990).  
 DR EMBL; M55334; AAA40974.1; -;  
 SQ SEQUENCE 18 AA; 1963 MW; 1D4F384151EF1DB CRC64;

Query Match 32.4%; Score 34; DB 11; Length 18;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GLNGCCRCGARGPE 18  
 I I I I I I I  
 DB 2 GWVGATRCQGRASE 15

RESULT 3  
 ID 000059 PRELIMINARY; PRT; 24 AA.  
 AC 000059;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE METALLOTHIONEIN-LIKE PROTEIN (PLANTA INDUCED RUST PROTEIN 11).  
 GN PG11.  
 OS Uromyces fabae.  
 OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Uredinales;  
 OC Pucciniaceae; Uromyces.  
 OX NCBI\_TaxID=55588;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=I2; TISSUE=HAUSTORIUM;  
 RA MEDLINE=97294919; PubMed=9150592;  
 RA Hahn M., Mengden K.;  
 RT "Characterization of in planta-induced rust genes isolated from a  
 haustorium-specific cDNA library.";  
 RL Mol. Plant Microbe Interact. 10:427-437(1997).  
 CC -|- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
 RESIDUES THAT BIND VARIOUS HEAVY METALS, SUCH AS COPPER, ZINC OR  
 CADMIUM. THESE PROTEINS ARE TRANSCRIPTIONALLY REGULATED BY BOTH  
 HEAVY METALS AND GLUCOCORTICOID. METALLOTHIONEINS PROTECT CELLS  
 AGAINST TOXIC CONCENTRATIONS OF HEAVY METALS AND ARE ACTIVE IN THE  
 PROTECTION AGAINST FREE OXYGEN RADICALS.  
 CC -|- DEVELOPMENTAL STAGE: HAUSTORIA AND RUST-INJECTED LEAVES. ALSO  
 OBSERVED, IN LOWER LEVELS, IN SPORES OR HYPHAE FORMED IN VITRO.  
 CC -|- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.  
 DR EMBL; U81791; AAB39879.1; -;  
 DR INTERPRO; IPR000006; -;  
 DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; 1.  
 KW Metal-binding; Copper; Chelation.

FT METAL 4 4 COPPER (BY SIMILARITY).  
 FT METAL 8 8 COPPER (BY SIMILARITY).  
 FT METAL 10 10 COPPER (BY SIMILARITY).  
 FT METAL 14 14 COPPER (BY SIMILARITY).  
 FT METAL 16 16 COPPER (BY SIMILARITY).  
 FT METAL 19 19 COPPER (BY SIMILARITY).  
 SQ SEQUENCE 24 AA; 2383 MW; 563C42F3C8A6F60E CRC64;

Query Match 29.0%; Score 30.5; DB 3; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 5.4e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 6 LNGC--CRCGA 14  
 : I I I I I  
 DB 1 MNPCCSSNCSGA 12

RESULT 4  
 P82706 PRELIMINARY; PRT; 16 AA.

ID P82706;  
 AC P82706;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE IMMUNE-INDUCED PROTEIN 1 (DIM-1).  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE.

RC STRAIN=OREGON-R; TISSUE=HEMOLYMPH;  
 RX MEDLINE=98409659; PubMed=9736738;  
 RA Uttenweiller-Joseph S., Moniatte M., Lagueux M., Van Dorselaer A.,  
 RA Hoffmann J.A., Bulet P.;  
 RT "Differential display of peptides induced during the immune response  
 of Drosophila: a matrix-assisted laser desorption ionization time-of-  
 flight mass spectrometry study.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:11342-11347(1998).  
 CC -|- TISSUE SPECIFICITY: HEMOLYMPH.  
 CC -|- INDUCTION: BY BACTERIAL INFECTION.  
 CC -|- MASS SPECTROMETRY: MW=1687.7; METHOD=MALDI.  
 KW Antibiotic; Insect immunity; Amidation.  
 FT MOD.RES 16 16  
 FT DISULFID 9 12  
 SQ SEQUENCE 16 AA; 1670 MW; 5F22356839CD06AF CRC64;

Query Match 28.1%; Score 29.5; DB 5; Length 16;  
 Best Local Similarity 50.0%; Pred. No. 5.4e+02;  
 Matches 6; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 6 LNGCCR-CGARG 16  
 : I I I I I I  
 DB 5 INGDCRVCNVHG 16

RESULT 5  
 Q14042 PRELIMINARY; PRT; 18 AA.

ID Q14042;  
 AC Q14042;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)  
 DE TYPE II PROCOLLAGEN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Strom C.M., Upholt W.B.;  
 RL Nucleic Acids Res. 11:1025-1037(1983).  
 DR EMBL; X00299; CAA25082.1; -;  
 SQ SEQUENCE 18 AA; 1595 MW; 17789165A92EEB0F CRC64;

Query Match 28.1%; Score 29.5; DB 4; Length 18;  
 Best Local Similarity 47.1%; Pred. No. 6e+02;  
 Matches 8; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 3 ASGLNGCC-RCGARGPE 18  
 I I I I I I I I I I  
 DB 2 ATGFGAAGRVGPGPE 18

RESULT 6  
 Q90790 ID Q90790 PRELIMINARY; PRT; 18 AA.  
 AC Q90790;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
 DE ALPHA-1 TYPE III COLLAGEN (FRAGMENT).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83135706; PubMed=6298201;  
 RA Yamada Y., Mudryj M., Sullivan M., de Crombrughe B.;  
 RT "Isolation and characterization of a genomic clone encoding chick  
 alpha 1 type III collagen.";  
 RL J. Biol. Chem. 258:2758-2761(1983).  
 DR EMBL; V00391; CAA23689.1;  
 SQ SEQUENCE 18 AA; 1732 MW; 439C34D5A9329EBC CRC64;

Query Match 26.7%; Score 28; DB 13; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 GARGP 17  
 Db 7 GARGP 11

RESULT 7  
 Q9WH14 ID Q9WH14 PRELIMINARY; PRT; 25 AA.  
 AC Q9WH14;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE GDP-MANNOSE DEHYDROGENASE (FRAGMENT).  
 GN ALGD.  
 OS Pseudomonas syringae (pv. syringae).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=321;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FF5;  
 RX MEDLINE=99287831; PubMed=10348861;  
 RA Fakhr M.K., Penaloza-Vazquez A., Chakrabarty A.M., Bender C.L.;  
 RT "Regulation of alginate biosynthesis in Pseudomonas syringae pv.  
 syringae.";  
 RL J. Bacteriol. 181:3478-3485(1999).  
 DR EMBL; AF131068; AAD31821.1;  
 FT NON\_TER 25  
 SQ SEQUENCE 25 AA; 2531 MW; DE2F6B5D4D19E1CA CRC64;

Query Match 26.2%; Score 27.5; DB 2; Length 25;  
 Best Local Similarity 71.4%; Pred. No. 1.5e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 8 GC-CRCG 13  
 Db 17 GCLCAGC 23

RESULT 8  
 Q28787 ID Q28787 PRELIMINARY; PRT; 24 AA.

Q28787;  
 AC 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE RELAXIN (FRAGMENT).  
 GN RLX.  
 OS Papio sp. (Baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Papio.  
 OX NCBI\_TaxID=61183;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BABA;  
 RA Evans B.B.A., Fu P., Tregear G.G.W.;  
 RL J. Mol. Endocrinol. 0:0-0(1993).  
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
 DR EMBL; Z27106; CAA81622.1;  
 DR HSSP; P04090; 6RLX.  
 DR INTERPRO; IPR000739;  
 DR PROSITE; PS00262; INSULIN; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 24 AA; 2691 MW; 560BC26A8BADBA14 CRC64;

Query Match 25.7%; Score 27; DB 6; Length 24;  
 Best Local Similarity 57.1%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 NGCCRCG 13  
 Db 8 NKCHIG 14

RESULT 9  
 Q28430 ID Q28430 PRELIMINARY; PRT; 24 AA.  
 AC Q28430;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE RELAXIN 2A (FRAGMENT).  
 GN RLX 2A.  
 OS Gorilla gorilla gorilla (Lowland gorilla).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.  
 OX NCBI\_TaxID=9595;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RG YK;  
 RA Evans B.B.A., Fu P., Tregear G.G.W.;  
 RL J. Mol. Endocrinol. 0:0-0(1993).  
 CC -1- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO  
 PRODUCE DILATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE  
 INVOLVED IN REMODELLING OF CONNECTIVE TISSUES DURING PREGNANCY,  
 PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.  
 CC -1- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO  
 DISULFIDE BONDS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CORPUS LUTEUM OF PREGNANCY  
 AND IN THE PLACENTA (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.  
 DR EMBL; Z27228; CAA81742.1;  
 DR HSSP; P04090; 6RLX.  
 DR INTERPRO; IPR000739;  
 DR PROSITE; PS00262; INSULIN; 1.  
 KW Insulin family; Hormone; Multigene family; Polymorphism.  
 FT NON\_TER 1  
 SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match 25.7%; Score 27; DB 6; Length 24;

```

Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 NGCCRCG 13
Db 8 NKCCCHVG 14

RESULT 10
Q28781
ID Q28781 PRELIMINARY; PRT; 24 AA.
AC Q28781;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE RELAXIN 1B (FRAGMENT).
GN RLX 1B.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RG YK 287;
RA Evans B. B. A., Fu P., Tregear G. G. W.;
RL J. Mol. Endocrinol. 0:0-0(1993).
CC -!- FUNCTION: RELAXIN IS AN OVARIAN HORMONE THAT ACTS WITH ESTROGEN TO
CC PRODUCE DILATION OF THE BIRTH CANAL IN MANY MAMMALS. MAY BE
CC INVOLVED IN REMODELLING OF CONNECTIVE TISSUES DURING PREGNANCY,
CC PROMOTING GROWTH OF PUBIC LIGAMENTS AND RIPENING OF THE CERVIX.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CORPUS LUTEUM OF PREGNANCY
CC AND IN THE PLACENTA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL: 227229; CAAB1743.1; -
DR INTERPRO: IPR000739; -
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Multigene family; Polymorphism.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match 25.7%; Score 27; DB 6; Length 24;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 NGCCRCG 13
Db 8 NKCCCHVG 14

RESULT 11
O89021
ID O89021 PRELIMINARY; PRT; 24 AA.
AC O89021;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MUSCULAR CHLORIDE CHANNEL 1 (FRAGMENT).
GN CIC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=SKELETAL MUSCLE;
RA Vullhorst D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ011106; CAA09489.1; -
FT NON_TER 1

```

---

```

SQ SEQUENCE 24 AA; 2502 MW; B1ABB4650DC0BFE4 CRC64;

Query Match 24.8%; Score 26; DB 11; Length 24;
Best Local Similarity 45.5%; Pred. No. 2.4e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAASGLNGCCR 11
Db 9 GISLGMVGCGR 19

RESULT 12
P82276
ID P82276 PRELIMINARY; PRT; 13 AA.
AC P82276;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE TUBULIN-ASSOCIATED PEPTIDE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RA Rossbacher J., Gartner W., Katinger H., Wagner L.;
RT "Haptoglobin mediated changes in cellular immune response.";
RL Submitted (JAN-2000) to the SWISS-PROT data bank.
CC -!- TISSUE SPECIFICITY: LEUKOCYTES AND HEPATOCYTES.
KW Microtubules.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1300 MW; 4FD05B1E478D62C7 CRC64;

Query Match 23.8%; Score 25; DB 4; Length 13;
Best Local Similarity 85.7%; Pred. No. 2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AASGLNG 8
Db 2 AASGLVG 8

RESULT 13
Q9UE38
ID Q9UE38 PRELIMINARY; PRT; 18 AA.
AC Q9UE38;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE COLLAGEN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89325561; PubMed=2753125;
RA Vikkula M., Peltonen L.;
RT "Structural analyses of the polymorphic area in type II collagen
RT gene.";
RL FEBS Lett. 250:171-174(1989).
DR EMBL: X16158; CAA34284.1; -
KW Collagen.
SQ SEQUENCE 18 AA; 1683 MW; 7724306D60CA21B4 CRC64;

Query Match 23.8%; Score 25; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 13 GARGPE 18  
| :|||  
Db 4 GEKGPE 9

## RESULT 14

Q9S1W0 PRELIMINARY; PRT; 19 AA.  
AC Q9S1W0  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE SMALL HYPOTHETICAL PROTEIN.  
GN SCJ4.13C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Saunders D.C., Harris D.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmid and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL109950; CAB52947.1; -  
SQ SEQUENCE 19 AA; 2028 MW; F643A294D0127DF6 CRC64;

Query Match 23.8%; Score 25; DB 2; Length 19;  
Best Local Similarity 60.0%; Pred. No. 2.8e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCCRC 12  
| :|:  
Db 12 GACQC 16

## RESULT 15

Q07101 PRELIMINARY; PRT; 21 AA.  
AC Q07101  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
DE PLACENTAL GROWTH FACTOR 2 (PLGF-2) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RX MEDLINE=93205407; PubMed=7681160;  
RA Maglione D., Guerriero V., Viglietto G., Ferraro M.G., Aprelikova O.,  
RA Altalo K., Del Vecchio S., Lei K.J., Chou J.Y., Persico M.G.;  
RT "Two alternative mRNAs coding for the angiogenic factor, placenta  
RT growth factor (PlGF), are transcribed from a single gene of chromosome  
RT 14."  
RL Oncogene 8:925-931(1993).  
DR EMBL; S57152; AAB25832.1; -  
KW Growth factor; Placenta.

FT NON\_TER 1 1  
NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2098 MW; F6A47DB8742B28DE CRC64;

Query Match 23.8%; Score 25; DB 4; Length 21;  
Best Local Similarity 62.5%; Pred. No. 3e+03;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 RCGARGPE 18  
| :|||  
Db 5 RAGGRGGE 12

Search completed: May 8, 2001, 15:16:07  
Job time: 533 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:07:10 ; Search time 113.31 Seconds  
(without alignments)  
9.081 Million cell updates/sec

Title: US-09-165-546A-12

Perfect score: 91

Sequence: 1 SKLLEFYLMFPATPWEA 18

Scoring table:

BLOSUM62

Gapol 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 165078

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_0401.\*

```

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	100.0	18	Y52439	Human tumour antig
2	54	59.3	10	Y06013	Human cancer antig
3	52	57.1	10	Y06004	Human cancer antig
4	52	57.1	10	Y06007	Human cancer antig
5	51	56.0	10	Y06010	Human cancer antig
6	44	48.4	9	Y06021	Human cancer antig
7	44	48.4	9	Y17845	NY-ESO-1 derived H
8	37	40.7	11	B12647	Protein kinase pep
9	37	40.7	15	Y66484	Human MAP kinase S
10	37	40.7	15	Y66485	Human MAP kinase S
11	37	40.7	15	Y66509	Human MAP kinase S

12	37	40.7	15	21	Y66510	Human MAP kinase S
13	36	39.6	9	21	Y79757	NY-ESO-1 derived p
14	32	35.2	17	15	R49308	Transferrin recept
15	31	34.1	16	18	W10317	Beta-subtype retin
16	30	33.0	9	21	Y79752	NY-ESO-1 derived p
17	30	33.0	10	20	Y06056	Human cancer antig
18	30	33.0	10	20	Y05990	Human cancer antig
19	30	33.0	12	18	W16119	Peptide containing
20	30	33.0	12	19	W68759	Peptide binding in
21	30	33.0	12	19	W58246	IL-IRTI binding pe
22	30	33.0	12	20	Y09890	Interleukin-1 type
23	30	33.0	12	21	B17671	IL-1 antagonist pe
24	30	33.0	14	14	R36921	Insulin-like growt
25	30	33.0	24	19	W68467	Amino acids 139-16
26	29	31.9	14	19	W59119	FMDV non-structura
27	29	31.9	19	19	W68469	Junction region of
28	29	31.9	24	20	Y27475	E. coli beta-subun
29	28	30.8	10	18	W16102	Peptide containing
30	28	30.8	10	19	W68740	Peptide binding in
31	28	30.8	10	19	W58219	IL-IRTI binding pe
32	28	30.8	10	20	Y09869	Interleukin-1 type
33	28	30.8	10	21	B17654	IL-1 antagonist pe
34	28	30.8	14	20	Y45370	Human secreted pro
35	28	30.8	21	20	W88700	Secreted protein e
36	27	29.7	9	21	Y82843	Lactadherin (BA-46
37	27	29.7	10	18	W16201	Peptide containing
38	27	29.7	10	18	W16023	Peptide containing
39	27	29.7	10	19	W68661	Random peptide seq
40	27	29.7	10	19	W68840	Peptide binding in
41	27	29.7	10	19	W58140	Interleukin-1 type
42	27	29.7	10	19	W58327	IL-IRTI binding pe
43	27	29.7	10	20	Y09971	Interleukin-1 type
44	27	29.7	10	20	Y09784	Interleukin-1 type
45	27	29.7	10	21	B17588	IL-1 antagonist pe

#### ALIGNMENTS

```

RESULT 1
Y52439
ID Y52439 standard; Protein; 18 AA.
XX
AC Y52439;
XX
DT 15-FEB-2000 (first entry)
XX
DE Human tumour antigen NY-ESO-1 peptide #12.
XX
KW Cancer; tumour; antigen; MHC; major histocompatibility complex; Class II;
KW T-cell; helper; stimulation; proliferation; treatment;
KW diagnosis; prevention; melanoma; breast cancer; ovarian cancer;
KW prostate cancer; hepatoma; thyroid cancer; bladder cancer; lung cancer;
KW lymphoma.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9953938-A1.
XX
PD 28-OCT-1999.
XX
PF 24-MAR-1999; 99WO-US06875.
XX
PR 17-APR-1998; 98US-0062422.
XX
PA 02-OCT-1998; 98US-0165546.
XX
(LUDW-) LUDWIG INST CANCER RES.
XX
PI Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
PI Gure A, Ritter G;
XX
DR WPI; 2000-038483/03.

```

XX Novel peptides which bind to MHC class I and MHC class II molecules,  
 PT useful for therapeutic and diagnostic purposes  
 XX  
 PS Claim 4; Page 22; 49pp; English.  
 XX  
 CC Peptides #8-#13 (Y52435-Y52440) are peptides derived from the human  
 CC tumour antigen, NY-ESO-1 (Y52430) which can bind to MHC(major  
 CC histocompatibility Class II HLA-DR53 molecules, thereby stimulating  
 CC proliferation of helper T-cells. cDNA encoding NY-ESO-1 was initially  
 CC isolated from an oesophagus squamous cell cancer cDNA library. Tissue  
 CC localisation studies revealed it to be expressed at high levels  
 CC in normal ovary and testis but not in normal colon, kidney, liver,  
 CC brain, oesophagus and skin. It was expressed in certain tumours and  
 CC tumour cell lines with some degree of frequency - these included  
 CC melanoma specimens and cell lines, and breast and bladder cancer  
 CC specimens, with expression in other tumour types being sporadic.  
 CC These NY-ESO-1-derived peptides may be used in methods and  
 CC compositions used for the treatment, diagnosis and prevention of  
 CC cancers (such as melanoma, breast cancer, prostate cancer, lung  
 CC cancer, hepatoma, ovarian cancer, thyroid cancer, bladder cancer,  
 CC or lymphoma) and to stimulate the proliferation of T cells.  
 XX  
 SQ Sequence 18 AA;

Query Match 100.0%; Score 91; DB 21; Length 18;  
 Best Local Similarity 100.0%; Pred. NO. 2.8e-09;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRLLEFYLAMPTPMEV 18  
 Db 1 srllefyiamptpmea 18  
 |||||

RESULT 2  
 Y06013  
 ID Y06013 standard; Peptide; 10 AA.

XX  
 AC Y06013;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.  
 DE  
 KW NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PA Rosenberg SA, Wang RF;  
 XX  
 PI WPI; 1999-277270/23.  
 XX  
 DR Cancer antigen NY ESO1/CAG-3  
 XX  
 PT Example 10; Page 42; 88pp; English.  
 XX

CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 26) corresponds to amino acid residues  
 CC 91-100 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 SQ Sequence 10 AA;

Query Match 59.3%; Score 54; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0036;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 YLAMPTPMP 16  
 Db 1 ylamptpmp 10  
 |||||

RESULT 3  
 Y06004  
 ID Y06004 standard; Peptide; 10 AA.

XX  
 AC Y06004;  
 XX  
 DT 16-AUG-1999 (first entry)  
 XX Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.  
 DE  
 KW NY ESO-1/CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09918206-A2.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 21-SEP-1998; 98WO-US19609.  
 XX  
 PR 08-OCT-1997; 97US-0061428.  
 XX  
 PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PA Rosenberg SA, Wang RF;  
 XX  
 PI WPI; 1999-277270/23.  
 XX  
 DR Cancer antigen NY ESO1/CAG-3  
 XX  
 PT Example 10; Page 42; 88pp; English.  
 XX

CC This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 17) corresponds to amino acid residues  
 CC 87-96 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune

CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.  
 XX  
 SQ Sequence 10 AA;

Query Match 57.1%; Score 52; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0079;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LLEFYLAMPF 12  
 Db 1 lleylamlpf 10

RESULT 4  
 Y06007  
 ID Y06007 standard; Peptide; 10 AA.

XX AC Y06007;

XX DT 16-AUG-1999 (first entry)

XX DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX OS Homo sapiens.

XX PN W09918206-A2.

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenberg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 42; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 20) corresponds to amino acid residues  
 CC 93-102 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,

CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX SQ Sequence 10 AA;

Query Match 57.1%; Score 52; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.0079;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 AMPFATPMEA 18  
 Db 1 ampfatpmea 10

RESULT 5

Y06010  
 ID Y06010 standard; Peptide; 10 AA.

XX AC Y06010;

XX DT 16-AUG-1999 (first entry)

XX DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX OS Homo sapiens.

XX PN W09918206-A2.

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenberg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX Cancer antigen NY ESO1/CAG-3

XX Example 10; Page 42; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y05988-Y06017) were identified.  
 CC The present peptide (ranked 23) corresponds to amino acid residues  
 CC 86-95 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX SQ Sequence 10 AA;

Query Match 56.0%; Score 51; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLLEFYLAM 11  
 Db 1 rlllefylamp 10  
 |||||

## RESULT 6

ID Y06021 standard; Peptide; 9 AA.

XX AC Y06021;

DT 16-AUG-1999 (first entry)

DE Human cancer antigen NY ESO-1/CAG-3 HLA peptide motif.

XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer peptide; antigen; human;  
 KW leukaemia; non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer;  
 KW metastasis; melanoma; adenocarcinoma; thymoma; colon cancer;  
 KW uterine cancer; breast cancer; prostate cancer; ovarian cancer;  
 KW cervical cancer; bladder cancer; kidney cancer; pancreatic cancer;  
 KW liver cancer; sarcoma; tumour; diagnosis; immunotherapy; therapy;  
 KW vaccine; human leukocyte antigen; HLA.

XX OS Homo sapiens.

XX PN WO9918206-A2.

XX PD 15-APR-1999.

XX PF 21-SEP-1998; 98WO-US19609.

XX PR 08-OCT-1997; 97US-0061428.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Rosenburg SA, Wang RF;

XX DR WPI; 1999-277270/23.

XX PT Cancer antigen NY ESO1/CAG-3

XX PS Example 10; Page 43; 88pp; English.

XX This peptide was identified as an HLA peptide motif following a  
 CC screen for epitopes from the coding region of human ESO-1/CAG-3  
 CC ORF1 (see X58599). 30 Epitopes (see Y06018-47) were identified.  
 CC The present peptide (ranked 4) corresponds to amino acid residues  
 CC 86-94 of CAG-1 ORF1 (see Y05965). CAG-1 is a new and potent  
 CC tumour antigen capable of eliciting an antigen specific immune  
 CC response by T cells. Cancer peptides (see Y05967-87) derived from  
 CC CAG-3, portions of CAG-3 and their variants, are useful as cancer  
 CC vaccines. A claimed method of preventing or inhibiting cancer  
 CC involves administering a cancer peptide, with or without an HLA  
 CC molecule. The cancer peptides form part of, or are derived  
 CC from, cancers such as primary or metastatic melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine  
 CC cancer, cervical cancer, bladder cancer, kidney cancer and  
 CC adenocarcinomas such as breast, prostate, ovarian, pancreatic and  
 CC thyroid cancers.

XX Sequence 9 AA;

Query Match 48.4%; Score 44; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLLEFYLAM 10  
 Db 1 rlllefylam 9  
 |||||

RESULT 7  
 Y78465  
 ID Y78465 standard; Peptide; 9 AA.  
 XX AC Y78465;  
 XX 10-MAY-2000 (first entry)  
 XX DE NY-ESO-1 derived HLA binding peptide #1.  
 XX Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;  
 KW HLA binding; human leukocyte antigen; cytolytic T cell; CTL;  
 KW cytostatic; melanoma; synovial sarcoma.  
 XX OS Homo sapiens.  
 XX PN WO200000824-A1.  
 XX PD 06-JAN-2000.  
 XX PF 25-JUN-1999; 99WO-US14493.  
 XX PR 26-JUN-1998; 98US-0105839.  
 XX PA (LUDW-) LUDWIG INST CANCER RES.  
 XX PI Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;  
 PI Chen Y, Gure A, Old LJ;  
 XX WPI; 2000-170933/15.  
 XX Determining the possible presence of breast, endometrial, colorectal,  
 PT lung, bladder or head-neck cancer -  
 XX Claim 34; Page 33; 40pp; English.

A method has been developed for determining the possible presence of a cancer, which is not melanoma or synovial sarcoma. The method comprises assaying a sample taken from the subject to determine the expression of an SSX gene, and determining the expression as a determination of the possible presence of cancer. Expression of SSX1 gene indicates possible presence of breast, endometrial, colorectal, lung, bladder or head-neck cancer. These cancers are also detected by SSX2 and SSX4 gene expression. SSX2 gene expression additionally indicates possible presence of lymphoma, renal cell cancer, glioma and prostate cancer. Expression of SSX4 gene also indicates possible presence of ovarian or stomach cancer. SSX5 gene expression indicates the same cancers as SSX1, except breast cancer. Determining expression of SSX gene can be used to monitor progress of melanoma or synovial sarcoma, which is not cancer. The SSX-derived peptide complex stimulates proliferation of cytolytic T cells. This is useful for treating cancer, especially melanoma. Y78464 to Y78468 represent specifically claimed HLA binding peptides for use in the method of the invention. 288452 to 288465 represent PCR primers used in the isolation of SSX genes in the exemplification of the present invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides derived from SSX proteins or NY-ESO-1, which are used in the exemplification of the present invention.

XX Sequence 9 AA;

Query Match 48.4%; Score 44; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLLEFYLAM 10  
 Db 1 rlllefylam 9  
 |||||

```

RESULT 8
B12647
ID B12647 standard; peptide; 11 AA.
XX
XX AC B12647;
XX
XX 13-NOV-2000 (first entry)
DT
XX
XX DE Protein kinase peptide ERK6 human.
DE
XX
XX KW Protein kinase; P38-gamma; crystallised P38 complex; structure;
KW protein co-ordinate data; inhibitor; identification.
XX
XX OS Homo sapiens.
XX
XX PN WO200036096-A1.
XX
XX PD 22-JUN-2000.
XX
XX PF 08-DEC-1999; 99WO-US29096.
XX
XX PR 16-DEC-1998; 98US-0112354.
XX
XX PR 03-NOV-1999; 99US-0163373.
XX
XX PA (VERT-) VERTEX PHARM INC.
XX
XX PI Bellon S, Bemis G, Wilson K, Fitzgibbon M;
XX
XX DR WPI; 2000-548618/50.
XX
XX PT Crystallized protein-kinase ligand complexes and their structural
XX coordinates useful for designing and identifying protein kinase
XX inhibitors.
XX
XX PS Claim 5; Page 33; 179pp; English.
XX
XX CC The present invention describes crystallised protein-kinase ligand
XX complexes (especially P38-ligand complexes) and their structure
XX coordinates (SCs). The key structural features of the proteins,
XX especially the shape of the substrate binding site, are useful in
XX methods for designing and/or identifying selective inhibitors of
XX protein kinases and in solving the structures of other proteins with
XX similar features. The SCs may be displayed on a computer for graphical
XX 3 dimensional representation of the structure and for computer aided
XX molecular design of new inhibitors. The SCs are based on the structure
XX of the phosphorylated P38-gamma complex which has been solved and which
XX reveals new structural information useful for understanding the
XX activated states of other, related kinase proteins. The present sequence
XX represents a protein kinase peptide sequence which is used in the
XX exemplification of the present invention.
XX
XX SQ Sequence 11 AA;

Query Match 40.7%; Score 37; DB 21; Length 11;
Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EFYLAMPF 12
DB 2 dfylvmpf 9

RESULT 9
Y66484
ID Y66484 standard; peptide; 15 AA.
XX
XX AC Y66484;
XX
XX DT 07-MAR-2000 (first entry)
XX

Human MAP kinase SAPK3 peptide fragment.
DE
XX
XX KW Protein kinase; inhibitor; extracellular matrix deposition; ECM; cancer;
KW tissue regeneration; tissue remodelling; wound healing; injury; fibrosis;
KW immunosuppression; diabetic nephropathy; and kidney damage; bone growth;
KW MAP kinase; osteoporosis; arthritis; atherosclerosis; cytokine; tumor.
XX
XX OS Homo sapiens.
XX
XX PN WO9958128-A1.
XX
XX PD 18-NOV-1999.
XX
XX PF 04-MAY-1999; 99WO-GB01385.
XX
XX PR 09-MAY-1998; 98GB-0009869.
XX
XX PA (MEDI-) MEDICAL RES COUNCIL.
XX
XX PI Cohen P, Goedert M, Evers PA;
XX
XX DR WPI; 2000-038954/03.
XX
XX PT Inhibiting particular protein kinases that have non-bulky residue in
XX the ATP binding site using pyridinyl-imidazole derivative, e.g. for
XX reducing deposition of extracellular matrix.
XX
XX PS Example 1; Fig 1; 93pp; English.
XX
XX CC The invention relates to inhibiting a protein kinase (I) that has Thr (or
XX a residue less bulky than Thr) at the position equivalent to Thr 106 in
XX SAPK2a/p38 by treatment with a pyridinyl-imidazole, or related inhibitor.
XX The (I) is not SAPK2a/p38 or SAPK2b/p38beta2. The inhibitors are used
XX therapeutically: (i) to treat extracellular matrix (ECM) deposition, to
XX encourage tissue regeneration and/or repair, tissue remodelling and
XX healing of wounds or injuries, also to reduce scar tissue formation
XX following a brain injury; (ii) to treat patients with (or at risk of
XX developing) end-stage organ failure, pathological ECM accumulation,
XX (renal) fibrosis, immunosuppression, diabetic nephropathy, and kidney
XX damage; (iii) to treat defects of bone growth or homeostasis (e.g.
XX osteoporosis), arthritis and atherosclerosis in which transforming growth
XX factorbeta (but not proinflammatory cytokines) are implicated, and (iv)
XX to treat epithelial, mesenchymal or glial cancers, solid tumors,
XX melanomas, small cell lung cancer, chronic myelogenous leukemia
XX (including purging bone marrow of affected cells) and T cell lymphoma.
XX Transgenic animals (or transfected cells) that include a protein kinase
XX in which the position equivalent to 106 is mutated are used to determine
XX the physiological role of the kinase and to identify its substrates, and
XX the protein kinase can be used to screen for drugs (or lead compounds).
XX The inhibitors are specific for the (I), i.e. they do not inhibit similar
XX protein kinases in which the position equivalent to Thr 106 is occupied
XX by Met or some other bulky residue.
XX
XX SQ Sequence 15 AA;

Query Match 40.7%; Score 37; DB 21; Length 15;
Best Local Similarity 75.0%; Pred. No. 4.7;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EFYLAMPF 12
DB 3 dfylvmpf 10

RESULT 10
Y66485
ID Y66485 standard; peptide; 15 AA.
XX
XX AC Y66485;
XX
XX DT 07-MAR-2000 (first entry)
XX

```

DE Human MAP kinase SAPK4 peptide fragment.  
 XX Protein kinase; inhibitor; extracellular matrix deposition; ECM; cancer;  
 KW tissue regeneration; tissue remodelling; wound healing; injury; fibrosis;  
 KW immunosuppression; diabetic nephropathy; and kidney damage; bone growth;  
 KW MAP kinase; osteoporosis; arthritis; atherosclerosis; cytokine; tumor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9958128-A1.  
 XX  
 PD 18-NOV-1999.  
 XX  
 PF 04-MAY-1999; 99WO-GB01385.  
 XX  
 PR 09-MAY-1998; 98GB-0009869.  
 XX  
 PA (MEDI-) MEDICAL RES COUNCIL.  
 XX  
 PI Cohen P, Goedert M, Evers PA;  
 XX  
 DR WPI; 2000-038954/03.  
 XX  
 XX Inhibiting particular protein kinases that have non-bulky residue in  
 PT the ATP binding site using pyridinyl-imidazole derivative, e.g. for  
 PT reducing deposition of extracellular matrix -  
 XX  
 PS Example 1; Fig 1; 93pp; English.  
 XX  
 CC The invention relates to inhibiting a protein kinase (I) that has Thr (or  
 CC a residue less bulky than Thr) at the position equivalent to Thr 106 in  
 CC SAPK2a/p38 by treatment with a pyridinyl-imidazole, or related inhibitor.  
 CC The (I) is not SAPK2a/p38 or SAPK2b/p38beta2. The inhibitors are used  
 CC therapeutically: (i) to treat extracellular matrix (ECM) deposition, to  
 CC encourage tissue regeneration and/or repair, tissue remodelling and  
 CC healing of wounds or injuries, also to reduce scar tissue formation  
 CC following a brain injury; (ii) to treat patients with (or at risk of  
 CC developing) end-stage organ failure, pathological ECM accumulation,  
 CC (renal) fibrosis, immunosuppression, diabetic nephropathy, and kidney  
 CC damage; (iii) to treat defects of bone growth or homeostasis (e.g.  
 CC osteoporosis), arthritis and atherosclerosis in which transforming growth  
 CC factorbeta (but not proinflammatory cytokines) are implicated, and (iv)  
 CC to treat epithelial, mesenchymal or glial cancers, solid tumors,  
 CC melanomas, small cell lung cancer, chronic myelogenous leukemia  
 CC (including purging bone marrow of affected cells) and T cell lymphoma.  
 CC Transgenic animals (or transfected cells) that include a protein kinase  
 CC in which the position equivalent to 106 is mutated are used to determine  
 CC the physiological role of the kinase and to identify its substrates, and  
 CC the protein kinase can be used to screen for drugs (or lead compounds).  
 CC The inhibitors are specific for the (I), i.e. they do not inhibit similar  
 CC protein kinases in which the position equivalent to Thr 106 is occupied  
 CC by Met or some other bulky residue.  
 XX  
 SQ Sequence 15 AA;

Query Match 40.7%; Score 37; DB 21; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 4.7;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYIAMPF 12  
 :||| |||  
 Db 3 dfylvmpf 10

RESULT 11  
 ID Y66509 standard; peptide; 15 AA.  
 XX  
 AC Y66509;  
 XX  
 DT 07-MAR-2000 (first entry)  
 XX

DE Human MAP kinase SAPK3 peptide fragment.

XX Protein kinase; inhibitor; extracellular matrix deposition; ECM; cancer;  
 KW tissue regeneration; tissue remodelling; wound healing; injury; fibrosis;  
 KW immunosuppression; diabetic nephropathy; and kidney damage; bone growth;  
 KW MAP kinase; osteoporosis; arthritis; atherosclerosis; cytokine; tumor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9958128-A1.  
 XX  
 PD 18-NOV-1999.  
 XX  
 PF 04-MAY-1999; 99WO-GB01385.  
 XX  
 PR 09-MAY-1998; 98GB-0009869.  
 XX  
 PA (MEDI-) MEDICAL RES COUNCIL.  
 XX  
 PI Cohen P, Goedert M, Evers PA;  
 XX  
 DR WPI; 2000-038954/03.  
 XX  
 XX Inhibiting particular protein kinases that have non-bulky residue in  
 PT the ATP binding site using pyridinyl-imidazole derivative, e.g. for  
 PT reducing deposition of extracellular matrix -  
 XX  
 PS Example 1; Fig 1; 93pp; English.

XX The invention relates to inhibiting a protein kinase (I) that has Thr (or  
 CC a residue less bulky than Thr) at the position equivalent to Thr 106 in  
 CC SAPK2a/p38 by treatment with a pyridinyl-imidazole, or related inhibitor.  
 CC The (I) is not SAPK2a/p38 or SAPK2b/p38beta2. The inhibitors are used  
 CC therapeutically: (i) to treat extracellular matrix (ECM) deposition, to  
 CC encourage tissue regeneration and/or repair, tissue remodelling and  
 CC healing of wounds or injuries, also to reduce scar tissue formation  
 CC following a brain injury; (ii) to treat patients with (or at risk of  
 CC developing) end-stage organ failure, pathological ECM accumulation,  
 CC (renal) fibrosis, immunosuppression, diabetic nephropathy, and kidney  
 CC damage; (iii) to treat defects of bone growth or homeostasis (e.g.  
 CC osteoporosis), arthritis and atherosclerosis in which transforming growth  
 CC factorbeta (but not proinflammatory cytokines) are implicated, and (iv)  
 CC to treat epithelial, mesenchymal or glial cancers, solid tumors,  
 CC melanomas, small cell lung cancer, chronic myelogenous leukemia  
 CC (including purging bone marrow of affected cells) and T cell lymphoma.  
 CC Transgenic animals (or transfected cells) that include a protein kinase  
 CC in which the position equivalent to 106 is mutated are used to determine  
 CC the physiological role of the kinase and to identify its substrates, and  
 CC the protein kinase can be used to screen for drugs (or lead compounds).  
 CC The inhibitors are specific for the (I), i.e. they do not inhibit similar  
 CC protein kinases in which the position equivalent to Thr 106 is occupied  
 CC by Met or some other bulky residue.

XX  
 SQ Sequence 15 AA;

Query Match 40.7%; Score 37; DB 21; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 4.7;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EYIAMPF 12  
 :||| |||  
 Db 3 dfylvmpf 10

RESULT 12  
 ID Y66510 standard; peptide; 15 AA.  
 XX  
 AC Y66510;  
 XX  
 DT 07-MAR-2000 (first entry)  
 XX

DE XX Human MAP kinase SAPK4 peptide fragment.

KW XX Protein kinase; inhibitor; extracellular matrix deposition; ECM; cancer;

KW XX tissue regeneration; tissue remodelling; wound healing; injury; fibrosis;

KW XX immunosuppression; diabetic nephropathy; and kidney damage; bone growth;

KW XX MAP kinase; osteoporosis; arthritis; atherosclerosis; cytokine; tumor.

OS XX Homo sapiens.

XX XX

PN XX WO958128-A1.

XX XX

PD XX 18-NOV-1999.

XX XX

PF XX 04-MAY-1999; 99WO-GB01385.

XX XX

PR XX 09-MAY-1998; 98GB-0009869.

XX XX

PA XX (MEDI-) MEDICAL RES COUNCIL.

XX XX

PI XX Cohen P, Goedert M, Evers PA;

XX XX

DR XX WPI; 2000-038954/03.

XX XX

PT XX Inhibiting particular protein kinases that have non-bulky residue in

PT XX the ATP binding site using pyridinyl-imidazole derivative, e.g. for

PT XX reducing deposition of extracellular matrix

XX XX

PS XX Example 1; Fig 1; 93pp; English.

XX XX

CC XX The invention relates to inhibiting a protein kinase (I) that has Thr (or

CC XX a residue less bulky than Thr) at the position equivalent to Thr 106 in

CC XX SAPK2a/p38 by treatment with a pyridinyl-imidazole, or related inhibitor.

CC XX The (I) is not SAPK2a/p38 or SAPK2b/p38beta2. The inhibitors are used

CC XX therapeutically: (i) to treat extracellular matrix (ECM) deposition, to

CC XX encourage tissue regeneration and/or repair, tissue remodelling and

CC XX healing of wounds or injuries, also to reduce scar tissue formation

CC XX following a brain injury; (ii) to treat patients with (or at risk of

CC XX developing) end-stage organ failure, pathological ECM accumulation,

CC XX (renal) fibrosis, immunosuppression, diabetic nephropathy, and kidney

CC XX damage; (iii) to treat defects of bone growth or homeostasis (e.g.

CC XX osteoporosis), arthritis and atherosclerosis in which transforming growth

CC XX factorbeta (but not proinflammatory cytokines) are implicated, and (iv)

CC XX to treat epithelial, mesenchymal or glial cancers, solid tumors,

CC XX melanomas, small cell lung cancer, chronic myelogenous leukemia

CC XX (including purging bone marrow of affected cells) and T cell lymphoma.

CC XX Transgenic animals (or transfected cells) that include a protein kinase

CC XX in which the position equivalent to 106 is mutated are used to determine

CC XX the physiological role of the kinase and to identify its substrates, and

CC XX the protein kinase can be used to screen for drugs (or lead compounds).

CC XX The inhibitors are specific for the (I), i.e. they do not inhibit similar

CC XX protein kinases in which the position equivalent to Thr 106 is occupied

CC XX by Met or some other bulky residue.

XX XX

SQ XX Sequence 15 AA;

Query Match 40.7%; Score 37; DB 21; Length 15;

Best Local Similarity 75.0%; Pred. No. 4.7;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 EFYLAAMPF 12

Db :||| |||

3 dfylyvmpf 10

RESULT 13

ID Y79757

XX Y79757 standard; Peptide; 9 AA.

XX AC Y79757;

XX XX

DT 10-MAY-2000 (first entry)

XX XX

DE XX

DE XX NY-ESO-1 derived peptide #13.

XX XX

KW XX Cancer; SSX family; SSX-1; SSX-2; SSX-3; SSX-4; SSX-5; NY-ESO-1;

KW XX HLA binding; human leukocyte antigen; cytolytic T cell; CTL;

KW XX cytostatic; melanoma; synovial sarcoma.

XX XX

OS XX Homo sapiens.

XX XX

PN XX WO200000824-A1.

XX XX

PD XX 06-JAN-2000.

XX XX

PF XX 25-JUN-1999; 99WO-US14493.

XX XX

PR XX 26-JUN-1998; 98US-0105839.

XX XX

PA XX (LUDW-) LUDWIG INST CANCER RES.

XX XX

PI XX Tureci O, Sahin U, Pfreundschuh M, Rammensee G, Stevanovic S;

PI XX Chen Y, Gure A, Old LJ;

XX XX

DR XX WPI; 2000-170933/15.

XX XX

PT XX Determining the possible presence of breast, endometrial, colorectal,

PT XX lung, bladder or head-neck cancer

XX XX

PS XX Example 13; Page 26; 40pp; English.

XX XX

CC XX A method has been developed for determining the possible presence of a

CC XX cancer, which is not melanoma or synovial sarcoma. The method comprises

CC XX assaying a sample taken from the subject to determine the expression of

CC XX an SSX gene, and determining the expression as a determination of the

CC XX possible presence of cancer. Expression of SSX1 gene indicates possible

CC XX presence of breast, endometrial, colorectal, lung, bladder or head-neck

CC XX cancer. These cancers are also detected by SSX2 and SSX4 gene expression.

CC XX SSX2 gene expression additionally indicates possible presence of

CC XX lymphoma, renal cell cancer, glioma and prostate cancer. Expression of

CC XX SSX4 gene also indicates possible presence of ovarian or stomach cancer.

CC XX SSX5 gene expression indicates the same cancers as SSX1, except breast

CC XX cancer. Determining expression of SSX gene can be used to monitor

CC XX progress of melanoma or synovial sarcoma, which is not cancer. The

CC XX SSX-derived peptide complex stimulates proliferation of cytolytic T

CC XX cells. This is useful for treating cancer, especially melanoma. Y78464

CC XX to Y78468 represent specifically claimed HLA binding peptides for use in

CC XX the method of the invention. Z88452 to Z88465 represent PCR primers used

CC XX in the isolation of SSX genes in the exemplification of the present

CC XX invention. Y78469 to Y78500, and Y79684 to Y79762 represent peptides

CC XX derived from SSX proteins or NY-ESO-1, which are used in the

CC XX exemplification of the present invention.

XX XX

SQ XX Sequence 9 AA;

Query Match 39.6%; Score 36; DB 21; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 FATPMEA 18

Db |||||

1 fatpmea 7

RESULT 14

R49308

ID R49308 standard; Protein; 17 AA.

XX XX

AC R49308;

XX XX

DT 16-SEP-1994 (first entry)

XX XX

DE Transferrin receptor position 680-696.

XX XX

KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;

KW major histocompatibility complex; class II; allotype; type I diabetes;  
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;  
 KW multiple sclerosis; transplant rejection; vaccine; MHC.

OS Homo sapiens.

PN WO9404171-A.

XX 03-MAR-1994.

XX 11-AUG-1993; 93WO-US07545.

XX 11-AUG-1992; 92US-0925460.

XX 15-JUN-1993; 93US-0925460.

PA (HARD ) HARVARD COLLEGE.

XX Chicx RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;

PI Vignali DA;

XX WPI; 1994-082825/10.

XX Novel immunomodulatory peptide(s) and nucleic acids - useful for  
 PT treatment of auto-immune diseases, transplant rejection and for  
 PT vaccination

XX Disclosure; Page 36; 139pp; English.

XX The sequences given in R49291-505 and R46981-7038 represent peptide  
 CC fragments of naturally-occurring immunomodulatory proteins. These  
 CC fragments are between 10-30 residues in length and bind to a human  
 CC major histocompatibility complex (MHC) class II allotype. These  
 CC peptides may be used for therapy of autoimmune diseases, such as  
 CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to  
 CC reduce transplant rejection. They may also be used for vaccination  
 CC providing an exclusively T-cell-mediated response, which can be  
 CC class I or class-II based, or both, depending on the length and  
 CC character of the immunogenic peptides.

XX Sequence 17 AA;

Query Match 35.2%; Score 32; DB 15; Length 17;

Best Local Similarity 26.7%; Pred. No. 39;

Matches 4; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 4 LEFYLAMPFATPMEA 18

Db :|::|:|:|:

2 veyflspyspkex 16

RESULT 15

W10317

ID W10317 standard; peptide; 16 AA.

XX AC W10317;

XX 13-MAR-1997 (first entry)

DE Beta-subtype retinoic acid receptor, residues 11-26.

XX Beta-subtype; retinoic acid; receptor; beta-RAR;  
 KW amino-terminus; antibody; detection; immunoprecipitation assay;  
 KW epitope; tumour.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc\_difference 16

FT /note= "Cys added to natural sequence for KLH  
 FT conjugation"

XX US585244-A.

XX

PD 17-DEC-1996.

XX 10-JUN-1994; 94US-0258851.

XX 10-JUN-1994; 94US-0258851.

XX (LIGA-) LIGAND PHARM INC.

PA Allegretto EA, Pike JW;

XX WPI; 1997-051173/05.

XX Antibody to gamma sub-type retinoid X receptor - useful for  
 PT detecting receptor, esp. in tumour samples

XX Example 1; Columns 9-10; 20pp; English.

XX The present peptide comprises residues 11-26 of the beta-subtype  
 CC retinoic acid receptor (beta-RAR), N-terminus. It can be  
 CC used to generate antibodies for the detection of beta-RAR, which  
 CC following an immunoprecipitation assay was found to be present in  
 CC HL60 (human promyelocytic carcinoma) cell, HeLa S3 (human cervical  
 CC carcinoma) cell, Hep G2 (human hepatoma) cell, MCF-7 (human breast  
 CC carcinoma) cell minus beta-E2, MCF-7 cell plus beta-E2, MCF-7  
 CC tumour plus beta-E2 and ME-180 (human cervical carcinoma) tumour  
 CC extracts at respective levels of not determined, 9, 5, not  
 CC determined, not determined, not determined and not determined  
 CC fmol/ml.

XX Sequence 16 AA;

QY Query Match 34.1%; Score 31; DB 18; Length 16;

Best Local Similarity 50.0%; Pred. No. 54;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RLLEFYFLAMP 11

Db :|::|:|:|:

4 gildfyasp 13

Search completed: May 8, 2001, 15:07:10

Job time: 226 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:08:21 ; Search time 68.24 Seconds  
(without alignments)  
18.127 Million cell updates/sec

Title: US-09-165-546A-12  
Perfect score: 91  
Sequence: 1 SRLLEFYLAMPTPMEA 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 4929

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	29.7	15	2 B45115	peptidylprolyl iso
2	26	28.6	13	2 PH1596	Ig H chain V-D-J r
3	25	27.5	24	2 T08160	S locus-linked pro
4	24	26.4	25	2 B36934	orf3 3' of madA -
5	23	25.3	13	2 PH1593	Ig H chain V-D-J r
6	23	25.3	15	2 B61457	alpha-glucosidase
7	23	25.3	18	2 A48550	hexon - canine ade
8	23	25.3	18	2 B49048	T-cell receptor be
9	23	25.3	20	2 S78760	ribosomal protein
10	23	25.3	20	2 PX0061	beta-N-acetylgluco
11	23	25.3	25	2 S78343	hypothetical prote
12	23	25.3	25	2 S39360	CDK inhibitor - mo
13	22	24.2	19	2 S32675	nitrogen fixation
14	22	24.2	23	2 A53631	H-transferrin AT
15	22	24.2	24	2 S69080	GNF receptor alph
16	22	24.2	24	2 S38766	petroglaucin - Pet
17	22	24.2	25	2 S21204	H-transferrin AT
18	22	24.2	25	2 S69139	bumetanide-binding
19	21	23.1	14	2 S09721	2S albumin small c
20	21	23.1	14	2 PA0104	protein QF200070 -
21	21	23.1	14	2 C48394	major fat-globule
22	21	23.1	15	1 LFECF	pne operon leader
23	21	23.1	16	2 PH1580	Ig H chain V-D-J r
24	21	23.1	16	2 S68730	bleomycin-binding
25	21	23.1	18	2 PH1368	Ig heavy chain DJ
26	21	23.1	18	2 A59137	protein Pil - gold
27	21	23.1	20	2 C56894	intracrystalline c
28	21	23.1	20	2 B56894	intracrystalline c
29	21	23.1	21	2 PH1730	Ig heavy chain V r

30	21	23.1	23	2 JQ0281	hypothetical 2.5K
31	21	23.1	23	2 S58624	hypothetical prote
32	21	23.1	24	2 S59908	endothelial cell-v
33	21	23.1	25	2 S12997	pancreatic endopep
34	21	23.1	25	2 S17683	hypothetical prote
35	20	22.0	11	2 I52980	glucocerebrosidase
36	20	22.0	12	2 S43957	Ig mu chain V regi
37	20	22.0	14	2 PH1597	Ig H chain V-D-J r
38	20	22.0	14	2 PH1601	Ig H chain V-D-J r
39	20	22.0	14	2 PH1608	Ig H chain V-D-J r
40	20	22.0	15	2 A54397	ubiquitin-carrier
41	20	22.0	15	2 PA0027	protein QAL00006 -
42	20	22.0	15	2 PH1610	Ig H chain V-D-J r
43	20	22.0	16	2 PH1588	Ig H chain V-D-J r
44	20	22.0	16	2 PH1604	Ig H chain V-D-J r
45	20	22.0	16	2 S24667	protein-tyrosine k

ALIGNMENTS

RESULT 1  
B45115  
peptidylprolyl isomerase (EC 5.2.1.8) FKBP51 - human (fragment)  
N:Alternate names: FK506-binding protein FKBP51; peptidylprolyl cis-trans isomerase F  
C:Species: Homo sapiens (man)  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: B45115  
R:Wiederrecht, G.; Hung, S.; Chan, H.K.; Marcy, A.; Martin, M.; Calaycay, J.; Boulton  
J. Biol. Chem. 267, 21753-21760, 1992  
A:Title: Characterization of high molecular weight FK-506 binding activities reveals  
A:Reference number: A45115; MUID:93016131  
A:Accession: B45115  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15 <WIE>  
A:Experimental source: JURKAT cells  
A>Note: sequence extracted from NCBI backbone (NCBIP:116748)  
C:Keywords: cis-trans-isomerase; cyclosporin A binding

Query Match 29.7%; Score 27; DB 2; Length 15;  
Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 YLAMPEA 13  
||| |  
Db 7 YLAPPYA 13

RESULT 2  
PH1596  
Ig H chain V-D-J region (wild-type clone 303) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
C:Accession: PH1596  
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.  
J. Exp. Med. 178, 317-329, 1993  
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m  
A:Reference number: PH1580; MUID:93301609  
A:Accession: PH1596  
A:Molecule type: DNA  
A:Residues: 1-13 <LEV>  
A:Experimental source: bone marrow pre-B lymphocyte  
C:Keywords: immunoglobulin

Query Match 28.6%; Score 26; DB 2; Length 13;  
Best Local Similarity 41.7%; Pred. No. 2e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLLEFYLAMPF 12  
: | | | | :

Db 2 ARRLGYYAMDY 13

RESULT 3

T08160

S locus-linked protein SLL1b - rape

C:Species: Brassica napus (rape)

C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 11-May-2000

C:Accession: T08160

R:Yu, K.; Schafer, U.; Glavin, T.L.; Goring, D.R.; Rothstein, S.J.

Plant Cell 8, 2369-2380, 1996

A:Title: Molecular characterization of the S locus in two self-incompatible Brassica napus

A:Reference number: Z16388; MUID:97143881

A:Accession: T08160

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-24 <YUK>

A:Cross-references: EMBL:U66192; NID:gl518109; PIDN:AAB49422.1; PID:gl518111

Query Match 27.5%; Score 25; DB 2; Length 24;  
Best Local Similarity 57.1%; Pred. No. 5.7e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 ERYLAMP 11  
:|:|:|:|

Db 9 QFYLYVP 15

RESULT 4

B36934

orf3 3' of mada - Thiobacillus versutus (fragment)

C:Species: Thiobacillus versutus

C>Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 03-Mar-1995

C:Accession: B36934

R:Huitema, F.; van Beunum, J.; van Driessche, G.; Duine, J.A.; Canters, G.W.

J. Bacteriol. 175, 6254-6259, 1993

A:Title: Cloning and sequencing of the gene coding for the large subunit of methylamine

A:Reference number: A36934; MUID:94012487

A:Accession: B36934

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-25 <HUI>

A:Experimental source: ATCC 25364T

A>Note: sequence extracted from NCBI backbone (NCBIN:138060, NCBIP:138062)

Query Match 26.4%; Score 24; DB 2; Length 25;  
Best Local Similarity 54.5%; Pred. No. 8.9e+02;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEFYLAMPPAT 14  
:|:|:|:|

Db 14 LRSFLAALPAT 24

RESULT 5

PH1593

Ig H chain V-D-J region (wild-type clone 144) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999

C:Accession: PH1593

R:Levinson, D.A.; Campos-Torres, J.; Leder, P.

J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice

A:Reference number: PH1580; MUID:93301609

A:Accession: PH1593

A:Molecule type: DNA

A:Residues: 1-13 <LEV>

A:Experimental source: bone marrow pre-B lymphocyte

C:Keywords: immunoglobulin

---

Query Match 25.3%; Score 23; DB 2; Length 13;  
Best Local Similarity 41.7%; Pred. No. 6.7e+02;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRLLEFYLAMPPF 12  
:|:|:|:|

Db 2 ARPLRHHYAMDY 13

RESULT 6

B61457

alpha-glucosidase (EC 3.2.1.20) - Tetrahymena pyriformis (strain W) (fragment)

C:Species: Tetrahymena pyriformis

C>Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 07-Dec-1999

C:Accession: B61457

R:Banno, Y.; Sasaki, N.; Yoshino, T.; Mochizuki, J.I.; Hirata, H.; Nozawa, Y.

J. Protozool. 36, 562-567, 1989

A:Title: A thermostable acid alpha-glucosidase from Tetrahymena thermophila: purification

A:Reference number: A61457; MUID:90095988

A:Accession: B61457

A:Molecule type: protein

A:Residues: 1-15 <BAN>

C:Genetics:

A:Genetic code: SGC5

C:Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; lysosome; mo

Query Match 25.3%; Score 23; DB 2; Length 15;  
Best Local Similarity 42.9%; Pred. No. 7.8e+02;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 PFATPME 17  
:|:|:|

Db 4 PFTPLQ 10

RESULT 7

A48550

hexon - canine adenovirus 1 (fragment)

C:Species: Mastadenovirus can1 (canine adenovirus 1)

C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 26-Aug-1999

C:Accession: A48550

R:Cai, F.; Weber, J.M.

Virus Genes 6, 307-312, 1992

A:Title: Nucleotide and deduced amino acid sequence of the canine adenovirus type 1 p

A:Reference number: A48550; MUID:93033182

A:Accession: A48550

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18 <CAI>

A:Cross-references: GB:M72715; NID:g210022; PIDN:AAA42528.1; PID:g210023

A>Note: sequence extracted from NCBI backbone (NCBIN:114644, NCBIP:114646)

C:Superfamily: adenovirus hexon protein

Query Match 25.3%; Score 23; DB 2; Length 18;  
Best Local Similarity 57.1%; Pred. No. 9.4e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 YLAMPFA 13  
:|:|:|

Db 6 YLRTFPS 12

RESULT 8

B49048

T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 30-May-1997

C:Accession: B49048

R:Sloud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.

Eur. J. Immunol. 22, 2413-2418, 1992

A:Title: Limited heterogeneity of T cell receptor variable region gene usage in juven

A:Reference number: A49048; MUID:92387250

A:Accession: B49048

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-18 <SIO>

A:Experimental source: patient EV, IL-2R+ synovial T-cells

A:Note: sequence extracted from NCBI backbone (NCBIP:113264)

C:Keywords: T-cell receptor

Query Match 25.3%; Score 23; DB 2; Length 18;  
Best Local Similarity 66.7%; Pred. No. 9.4e+02;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 FYLAMP 11

Db 4 FYLCAP 9

RESULT 9

S78760

C:Species: Bos primigenius taurus (cattle)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: S78760

R:Graack, H.R.

submitted to the Protein Sequence Database, July 1999

A:Reference number: S78760

A:Accession: S78760

A:Molecule type: protein

A:Residues: 1-11;12-20 <GRA>

C:Keywords: mitochondrion

Query Match 25.3%; Score 23; DB 2; Length 20;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 RLLEFY 7

Db 11 RLAEY 16

RESULT 10

PX0061

beta-N-acetylglucosaminoglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 03-Mar-1995

C:Accession: PX0061

R:Kawano, J.; Oinuma, T.; Nakayama, T.; Suganuma, T.

J. Biochem. 111, 568-572, 1992

A:Title: Characterization of beta-1,4-galactosyltransferase purified from rat liver micr

A:Reference number: PX0061; MUID:92348341

A:Accession: PX0061

A:Molecule type: protein

A:Residues: 1-20 <KAW>

A:Experimental source: liver microsomes

C:Comment: This enzyme catalyzes the transfer of galactose from an activated UDP-galacto

C:Comment: The complex of this enzyme and alpha-lactalbumin is lactose synthase (EC 2.4.

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 25.3%; Score 23; DB 2; Length 20;  
Best Local Similarity 60.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 9 AMPFATPMEA 18

Db 1 AMPGATLQRA 10

RESULT 11

S78343

hypothetical protein 25 - Odontella sinensis chloroplast

C:Species: chloroplast Odontella sinensis

C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 24-Apr-1998

C:Accession: S78343

R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.

Plant Mol. Biol. Rep. 13, 336-342, 1995

A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sine

A:Reference number: S78238

A:Accession: S78343

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-25 <KOW>

A:Cross-references: EMBL:267753; NID:g1185127; PID:e211881; PID:g1185233

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995

C:Genetics:

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match 25.3%; Score 23; DB 2; Length 25;

Best Local Similarity 42.9%; Pred. No. 1.3e+03;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 LEFYLAM 10

Db 4 ISFYIAL 10

RESULT 12

S39360

CDK inhibitor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 19-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 10-Nov-1995

C:Accession: S39360

R:Gu, Y.; Turck, C.W.; Morgan, D.O.

Nature 366, 707-710, 1993

A:Title: Inhibition of CDK2 activity in vivo by an associated 20K regulatory subunit.

A:Reference number: S39360; MUID:94081957

A:Accession: S39360

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-25 <GUY>

Query Match 25.3%; Score 23; DB 2; Length 25;

Best Local Similarity 30.8%; Pred. No. 1.3e+03;

Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LLEFYLAMPFATP 15

Db 4 LTDFVHSKVYLS 16

RESULT 13

S32675

nitrogen fixation protein nifB - Anabaena variabilis (fragment)

C:Species: Anabaena variabilis

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000

C:Accession: S32675

R:Monnerjahn, U.; Boehme, H.

submitted to the EMBL Data Library, December 1992

A:Description: Cloning and expression in E. coli of the Anabaena.

A:Reference number: S32675

A:Accession: S32675

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-19 <MON>

A:Cross-references: EMBL:X69898; NID:g296503; PIDN:CAA49521.1; PID:g296504

C:Genetics:

A:Gene: nifB

C:Superfamily: Rhizobium nifB protein

Query Match 24.2%; Score 22; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LEFY 7  
| | | |  
Db 10 LEFY 13

RESULT 14  
A53631  
H+-transporting ATP synthase (EC 3.6.1.34) beta chain - thermophilic bacterium PS-3 (fra  
C:Species: thermophilic bacterium PS-3  
C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 17-Mar-1999  
C:Accession: A53631  
R:Paik, S.R.; Jault, J.M.; Allison, W.S.  
Biochemistry 33, 126-133, 1994  
A:Title: Inhibition and inactivation of the F-1 adenosinetriphosphatase from Bacillus ps  
A:Reference number: A53631; MUID:94114474  
A:Accession: A53631  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-23 <PAI>  
C:Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase al  
C:Keywords: ATP biosynthesis; hydrolase

Query Match 24.2%; Score 22; DB 2; Length 23;  
Best Local Similarity 33.3%; Pred. No. 1.8e+03;  
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 LEFY LAMPF 12  
::: |  
Db 2 IQFFLSONF 10

RESULT 15  
S69080  
GNF receptor alpha chain, precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 12-Feb-1998 #sequence\_revision 31-Jan-2000 #text\_change 31-Jan-2000  
C:Accession: S69080  
R:Treanor, J.J.S.; Goodman, L.; de Sauvage, F.; Stone, D.M.; Poulsen, K.T.; Beck, C.D.;  
vies, A.M.; Asai, N.; Takahashi, M.; Vandlen, R.; Henderson, C.E.; Rosenthal, A.  
Nature 382, 80-83, 1996  
A:Title: Characterization of a multicomponent receptor for GDNF.  
A:Reference number: S69080; MUID:96273032  
A:Accession: S69080  
A:Molecule type: mRNA  
A:Residues: 1-24 <TR>  
F:1-24/Domain: signal sequence #status predicted <SIG>

Query Match 24.2%; Score 22; DB 2; Length 24;  
Best Local Similarity 44.4%; Pred. No. 1.9e+03;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LLEFY LAMP 11  
| | |  
Db 3 LATLYFALP 11

Search completed: May 8, 2001, 15:08:22  
Job time: 293 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:08 ; Search time 40.07 Seconds  
(without alignments)  
15.388 Million cell updates/sec

Title: US-09-165-546A-12

Perfect score: 91

Sequence: 1 SRLLEFLAMPFATPWEA 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	24	26.4	14	1 ECDC_LYMDI	P80940 Lymantria d
2	23	25.3	18	1 HEX_ADECU	P35985 canine aden
3	23	25.3	25	1 YCX8_ODOSI	P49834 odontella s
4	22	24.2	14	1 MARI_ALTPSP	P29399 altermonas
5	22	24.2	19	1 ADC_CLOPA	P81336 clostridium
6	22	24.2	25	1 ATP0_SPIOL	P80082 spinacia ol
7	21	23.1	15	1 LPE_ECOLI	P03057 escherichia
8	21	23.1	15	1 UN01_PINPS	P81106 pinus pinas
9	21	23.1	22	1 ATP6_COTJA	P50681 coturnix co
10	21	23.1	25	1 BGBP_PENVA	P81182 penaeus van
11	20	22.0	13	1 CRBL_ICASP	P17237 icaria sp.
12	20	22.0	18	1 DRPH_UCAPU	P08871 uca pugilati
13	20	22.0	20	1 SODE_PASPI	P81527 pasteurella
14	20	22.0	25	1 SMBP_RAT	P80968 rattus norv
15	20	22.0	25	1 YIGI_SALTY	P40725 salmonella
16	19	20.9	9	1 MGMT_BOVIN	P29177 bos taurus
17	19	20.9	10	1 COXQ_SHEEP	P80337 ovis aries
18	19	20.9	12	1 PA2B_VIPBO	P31859 vipera beru
19	19	20.9	13	1 LMT4_LOCMI	P41490 locusta mig
20	19	20.9	15	1 MCRA_MEFFE	P22948 methanosarc
21	19	20.9	16	1 ANP8_ELEGR	P11921 eleginus gr
22	19	20.9	19	1 DHAB_COMTE	P80704 comamonas t
23	19	20.9	20	1 DPP4_BOVIN	P81425 bos taurus
24	19	20.9	21	1 YD90_HABIN	P45194 haemophilus
25	19	20.9	25	1 ANT3_MESAU	P81050 mesocricetu
26	18	19.8	8	1 UPAA_HUMAN	P30096 homo sapien
27	18	19.8	10	1 NS1_MYCTU	P81135 mycobacteri
28	18	19.8	11	1 PQOC_PSEFL	P55173 pseudomonas
29	18	19.8	17	1 ATPL_PAVLU	P28529 pavlova lut
30	18	19.8	19	1 ANP7_ELEGR	P11920 eleginus gr
31	18	19.8	21	1 ODP2_SOLTU	P81421 solanum tub
32	18	19.8	22	1 LANM_STRMU	P80666 streptococc
33	18	19.8	22	1 ODPX_BOVIN	P22439 bos taurus

34 18 19.8 22 1 VGLG\_RABVA P15199 rabies viru  
35 18 19.8 24 1 DNAJ\_STRAG P35694 streptococc  
36 18 19.8 24 1 HEMO\_LINRE P23544 lingula ree  
37 18 19.8 24 1 PQQA\_PSEFL P55171 pseudomonas  
38 18 19.8 25 1 NP4\_HUMAN P18078 homo sapien  
39 17 18.7 8 1 ALL7\_CARMA P81809 carcinus ma  
40 17 18.7 10 1 ANG1\_BOTJA Q10581 bothrops ja  
41 17 18.7 10 1 LPK3\_LOCMI P41488 locusta mig  
42 17 18.7 11 1 LPW\_THETH P05624 thermus aqu  
43 17 18.7 13 1 YPNP\_PHOLU P41122 photorhabdu  
44 17 18.7 14 1 HY14\_PIG P01155 sus scrofa  
45 17 18.7 15 1 ACEA\_ACICA P28467 acinetobact

#### ALIGNMENTS

RESULT 1  
ECDC\_LYMDI STANDARD; PRT; 14 AA.  
AC P80940;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE TESTIS ECDYSIOTROPIN PEPTIDE C (TE).  
OS Lymantria dispar (Gypsy moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Noctuoidea; Lymantriidae; Lymantria.  
OX NCBI\_Taxid=13123;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=97387807; PubMed=9243792;  
RA Loeb M.J.; Wagner R.M.; Woods C.W.; Gelman D.G.; Harrison D.;  
RA Bell R.A.;  
RT "Naturally occurring analogs of Lymantria testis ecdysiotropin, a gonadotropin isolated from brains of Lymantria dispar pupae.";  
RL Arch. Insect Biochem. Physiol. 36:37-50(1997).  
CC -!- FUNCTION: STIMULATES SYNTHESIS OF ECDYSTEROID IN THE TESTES OF LARVAE AND PUPAE.  
SQ SEQUENCE 14 AA; 1553 MW; 17F479531A685CBB CRC64;

Query Match 26.4%; Score 24; DB 1; Length 14;  
Best Local Similarity 57.1%; Pred. No. 3.5e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 YLAMPFA 13  
| :|||  
Db 7 YTPLPFA 13

RESULT 2  
HEX\_ADECU STANDARD; PRT; 18 AA.  
ID HEX\_ADECU  
AC P35985;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HEXON PROTEIN (LATE PROTEIN 2) (FRAGMENT).  
GN PII.  
OS Canine adenovirus type 1 (strain Utrecht).  
OC Viruses; GSDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_Taxid=36364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93033182; PubMed=1413543;  
RA Cai F.; Weber J.M.;  
RT "Nucleotide and deduced amino acid sequence of the canine adenovirus type 1 proteinase.";  
RL Virus Genes 6:307-312(1992).  
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE STRUCTURAL PROTEINS IN THE

```

CC VIRAL COAT AND IS SYNTHESIZED DURING LATE INFECTION.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M72715; AAA42528.1; -
CC PIR: A48550; A48550.
CC HSSP: P03277; 1DHX.
CC Coat protein; Hexon protein; Late protein.
CC NON_TER 1
CC SEQUENCE 18 AA; 1938 MW; B9E2AE307A420962 CRC64;

Query Match 25.3%; Score 23; DB 1; Length 18;
Best Local Similarity 57.1%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 YLAMPFA 13
   ||||
Db 6 YLTPFS 12

RESULT 3
YCX8_ODOSI STANDARD; PRT; 25 AA.
AC F49834;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 3.1 KDA PROTEIN IN PSBJ-TENE INTERGENIC REGION (ORF25).
OS Odontella sinensis.
OG Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscales; Eupodiscaeae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RL Odontella sinensis."
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z67753; CAA91716.1; -
CC Chloroplast; Hypothetical protein.
CC SEQUENCE 25 AA; 3083 MW; 9F50BB79484EA531 CRC64;

Query Match 25.3%; Score 23; DB 1; Length 25;
Best Local Similarity 42.9%; Pred. No. 9.1e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 LEFYLAM 10
   :||:
Db 4 ISFYIAL 10

RESULT 4
MARE_ALTSP STANDARD; PRT; 14 AA.

```

```

AC P29399;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE MARINOSTATINS C-2, C-1, AND D.
OS Alteromonas sp. (strain B-10-31).
OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC Alteromonas.
OX NCBI_TaxID=29456;
RN [1]
RP SEQUENCE, AND ACTIVE SITE.
RX MEDLINE=92176155; PubMed=1794974;
RA Takano R., Imada C., Kamei K., Hara S.;
RT "The reactive site of marinostatins, a proteinase inhibitor from
RL marine Alteromonas sp. B-10-31."
RL J. Biochem. 110:856-858(1991).
CC -1- FUNCTION: INHIBITS SUBTILISIN, CHYMOTRYPSIN, AND ELASTASE, BUT
CC NOT TRYPSIN.
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF MARINOSTATIN C-2.
KW Serine protease inhibitor.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT ACT_SITE 6 7 REACTIVE BOND.
FT VARIANT 1 2 MISSING (IN MARINOSTATIN C-1).
FT VARIANT 1 3 MISSING (IN MARINOSTATIN D).
SQ SEQUENCE 14 AA; 1644 MW; 6E7CEE92EF32E44 CRC64;

Query Match 24.2%; Score 22; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 PFAT 14
   ||||
Db 2 PFAT 5

RESULT 5
ADC_CLOPA STANDARD; PRT; 19 AA.
ID ADC_CLOPA
AC P81336;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ACETOACETATE DECARBOXYLASE (EC 4.1.1.4) (ADC) (CP 28/CP 29)
DE (FRAGMENT).
GN ADC.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RX STRAIN=W5;
RX MEDLINE=96291870; PubMed=9629918;
RA Flensburg R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5."
RL Electrophoresis 19:802-806(1998)
CC -1- CATALYTIC ACTIVITY: ACETOACETATE + H(+) = ACETONE + CO(2).
CC -1- SUBUNIT: HOMODODECAMER (BY SIMILARITY).
KW Lyase; Decarboxylase.
FT NON_TER 19
FT SEQUENCE 19 AA; 2078 MW; 6ACDAA91103AEC31 CRC64;

Query Match 24.2%; Score 22; DB 1; Length 19;
Best Local Similarity 37.5%; Pred. No. 1e+03;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 LAMPFATP 15
   :||:
Db 10 ISMPLTAP 17

```

```

RESULT 6
ATP0_SPTOL STANDARD; PRT; 25 AA.
AC P80082;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ATP SYNTHASE ALPHA CHAIN, MITOCHONDRIAL (EC 3.6.1.34)
DE (FRAGMENT).
GN ATPA.
OS Spinacia oleracea (Spinach).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN-CV. MEDANIA; TISSUE-Leaf mesophyll;
RX MEDLINE=92209531; PubMed=1313368;
RA Hanasur B., Glaser E.;
RT "Plant mitochondrial FOF1 ATP synthase. Identification of the
RT individual subunits and properties of the purified spinach leaf
RT mitochondrial ATP synthase.";
RL Eur. J. Biochem. 205:409-416(1992).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE. THE ALPHA CHAIN IS A REGULATORY
CC SUBUNIT.
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR PIR: S21204; S21204.
DR InterPro: IPR000194; -.
DR PROSITE: PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW ATP synthesis; CF(1); Hydrogen ion transport;
KW Hydrolase; ATP-binding; Mitochondrion.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2904 MW; 1B1486EBBD1A650D CRC64;

Query Match 24.2%; Score 22; DB 1; Length 25;
Best Local Similarity 57.1%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRLLEFY 7
Db 16 SRISNFY 22

RESULT 7
LPF_ECOLI STANDARD; PRT; 15 AA.
AC P03057;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHE LEADER PEPTIDE (ATTENUATOR PEPTIDE).
DE PHYL OR PHEAE.
GN ESCHERICHIA COLI.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=79033820; PubMed=360214;
RA Zurawski G., Brown K., Killingly D., Yanofsky C.;
RT "Nucleotide sequence of the leader region of the phenylalanine operon
RT of Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 75:4271-4275(1978).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91072346; PubMed=2254312;
RA Gavini N., Davidson B.E.;
RT "pheA mutants of Escherichia coli have a defective pheA attenuator.";
RL J. Biol. Chem. 265:21532-21535(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS
CC OF PHENYLALANINE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00314; CAA23600.1; -
DR EMBL: M10431; AAA24329.1; -
DR EMBL: M58024; AAA62783.1; -
DR EMBL: AE000346; AAC75647.1; -
DR PIR: A03593; LFECEP.
DR PIR: B36494; B36494.
DR EcoGene: EG11271; pheL.
KW Leader peptide.
SQ SEQUENCE 15 AA; 1924 MW; CFE14AE3BFF935E0 CRC64;

Query Match 23.1%; Score 21; DB 1; Length 15;
Best Local Similarity 33.3%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 4 LEFYLAMPTATP 15
Db 4 IPFFFAFFFTFP 15

RESULT 8
UNOL_PINPS STANDARD; PRT; 15 AA.
AC P81106;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (S1247/S1248) (N150/N151)
DE (FRAGMENT).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RA Plomion C., Costa P., Bahrman N., Frigerio J.M.;
RT "Genetic analysis of needle proteins in maritime pine. 1. Mapping
RT dominant and codominant protein markers assayed on diploid tissue, in
RT a haploid-based genetic map.";
RL Silvae Genetica 46:161-165(1997).
RN [2]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Plomion C., Bauw G., Dubos C., Bahrman N., Kremer A.,

```

RESULT	10
BGBP_PENVA	
ID	BGBP_PENVA STANDARD; PRT; 25 AA.
AC	P81182;
DT	15-JUL-1998 (Rel. 36, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	15-JUL-1998 (Rel. 36, Last annotation update)
DE	BETA 1,3-D-GLUCAN BINDING PROTEIN (BGBP) (FRAGMENT).
OS	Penaeus vannamei (Penaeid shrimp) (European white shrimp).
OC	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC	Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC	Penaeidae; Penaeus.
OX	NCBI_TaxID=6689;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=97293426; PubMed=9149399;
RA	Vargas-Albores F., Jimenez-Vega F., Yepiz-Plascencia G.M.;
RT	"Purification and comparison of beta-1,3-glucan binding protein from
RT	white shrimp (Penaeus vannamei).";
RL	Comp. Biochem. Physiol. 116B:453-458(1997).
CC	-I- FUNCTION: INVOLVED IN THE RECOGNITION OF INVADING MICRO-ORGANISMS
CC	-I- SUBUNIT: MONOMER.
NON_TER	25 25
FT	SEQUENCE 25 AA; 2743 MW; 747DD77EDC0FB06F CRC64;
SQ	
Query Match	23.1%; Score 21; DB 1; Length 25;
Best Local Similarity	44.4%; Pred. No. 2e+03;
Matches	4; Conservative 1; Mismatches 4; Indels 0; Gaps
QY	4 LEFLAMPF 12     :
Db	14 LRFNMKTPF 22
RESULT	11
CRBL_ICASP	
ID	CRBL_ICASP STANDARD; PRT; 13 AA.
AC	P17237;
DT	01-AUG-1990 (Rel. 15, Created)
DT	01-AUG-1990 (Rel. 15, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last annotation update)
DE	CHEMOTACTIC PEPTIDE (I-Cp).
OS	Icaria sp. (Ropalidian wasp).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Hymenoptera; Aculeata;
OC	Vespoidea; Vespidae; Polistinae; Icaria.
OX	NCBI_TaxID=7495;
RN	[1]
RP	SEQUENCE.
RT	TISSUE=Venom;
RA	Yasuhara T., Itokawa H., Suzuki N., Nakamura H., Nakajima T.;
RL	(in) Izumiya N. (eds.);
RL	Peptide chemistry 1984, pp.177-182, Protein Research Foundation,
RL	Osaka (1985).
CC	-I- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC	OF NEUTROPHILS.
KW	MAST cell degranulation; Chemotaxis; Venom; Amidation.
FT	MOD_RES 13 13 AMIDATION.
SQ	SEQUENCE 13 AA; 1353 MW; 348DBC7AA30A3768 CRC64;

```

Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps
0;

QY 10 MPFATPM 16
      :||| | :
Db 2 VPFLGPL 8
      |
RESULT 12

```

DRPH\_UCAPU STANDARD; PRT; 18 AA.  
AC P08871.  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE PIGMENT-DISPERSING HORMONE (PDH) (LIGHT ADAPTING DISTAL RETINAL  
DE PIGMENT HORMONE) (DRPH).  
OS Uca pugilator (Atlantic sand fiddler crab) (Celuca pugilator).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Ocypodoidea; Ocypodidae; Celuca.  
OX NCBI\_TaxID=6772;  
RN [1]  
RP SEQUENCE.  
RA Rao K.R., Richm J.P., Zahnow C.A., Kleinholz L.H., Tarr G.E.,  
RA Johnson L., Norton S., Landau M., Semmes O.J., Sattelberg R.M.,  
RA Jorenby W.H., Hintz M.F.;  
RT "Characterization of a pigment-dispersing hormone in eyestalks of the  
RT fiddler crab Uca pugilator.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:5319-5322(1985).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE=93230895; PubMed=8472537;  
RA Loehr J., Klein J., Webster S.G., Dirksen H.;  
RT "Quantification, immunoaffinity purification and sequence analysis of  
RT a pigment-dispersing hormone of the shore crab, Carcinus maenas  
RT (L.).";  
RL Comp. Biochem. Physiol. 104B:699-706(1993).  
CC -1- FUNCTION: CAUSES THE MIGRATION OF THE DISTAL RETINAL PIGMENT  
CC INTO THE PROXIMAL END OF THE PIGMENT CHROMATOPHORE CELLS AND  
CC THIS DECREASES THE AMOUNT OF LIGHT ENTERING THE RETINULAS.  
CC -1- SIMILARITY: TO THE PDH OF OTHER ARTHROPODS.  
DR PIR: A25144; DRUFPD.  
KW Hormone; Amidation.  
FT DOMAIN 6 9 IMPORTANT FOR DRPH ACTIVITY.  
FT MOD\_RES 18 18 AMIDATION.  
SQ SEQUENCE 18 AA; 1928 MW; 25D5CE8D016F544E CRC64;  
Query Match 22.0%; Score 20; DB 1; Length 18;  
Best Local Similarity 36.4%; Pred. No. 2.1e+03;  
Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 1 SRLLEFFYAMP 11  
| | | | |  
DB 2 SELINSILGLP 12  
| | | | |  
RESULT 13  
SODF\_PASPI STANDARD; PRT; 20 AA.  
AC P81527; 20 AA.  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE SUPEROXIDE DISMUTASE [FE] (EC 1.15.1.1) (FRAGMENT).  
GN SODB.  
OS Pasteurella piscicida (Photobacterium damsela (subsp. piscicida)).  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;  
OC Photobacterium.  
OX NCBI\_TaxID=38294;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=WT1415;  
RX MEDLINE=99173752; PubMed=10075430;  
RA Barnes A.C., Balebona M.C., Horne M.T., Ellis A.E.;  
RT "Superoxide dismutase and catalase in Photobacterium damsela subsp.  
RT piscicida and their roles in resistance to reactive oxygen species.";  
RL Microbiology 145:483-494(1999).  
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.  
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).  
CC

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.  
DR InterPro; IPR001189; -;  
DR Pfam; PF00081; sodfe; 1.  
KW Oxidoreductase; Iron; Periplasmic.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2153 MW; A8D31FDAE8553B6D CRC64;  
Query Match 22.0%; Score 20; DB 1; Length 20;  
Best Local Similarity 60.0%; Pred. No. 2.3e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 9 AMPFA 13  
| | | | |  
DB 6 ALPYA 10  
| | | | |  
RESULT 14  
SMBP\_RAT STANDARD; PRT; 25 AA.  
AC P80568; 25 AA.  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE SM-11044 BINDING PROTEIN (FRAGMENTS).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=WISTAR;  
RX MEDLINE=97407910; PubMed=9261134;  
RA Sugasawa T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,  
RA Morooka S., Strosberg A.D.;  
RT "Characterization of a novel iodoctyanopindolol and SM-11044 binding  
RT protein, which may mediate relaxation of depolarized rat colon  
RT tonus.";  
RL J. Biol. Chem. 272:21244-21252(1997).  
CC -1- FUNCTION: MAY MEDIATE RELAXATION OF DEPOLARIZED COLON TONUS.  
CC IT BINDS IODOCTYANOPINDOLOL AND SM-11044.  
CC -1- MISCELLANEOUS: THE ORDER OF THE FRAGMENTS IS UNKNOWN.  
FT NON\_TER 1 1  
FT UNSURE 6 6 OR Y.  
FT NON\_CONS 18 19  
FT NON\_TER 25 25  
SQ SEQUENCE 25 AA; 3177 MW; D14F0CB9B778C2CB CRC64;  
Query Match 22.0%; Score 20; DB 1; Length 25;  
Best Local Similarity 42.9%; Pred. No. 2.9e+03;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SRLLEFFY 7  
| | | | |  
DB 11 ARYFQFY 17  
| | | | |  
RESULT 15  
YIGI\_SALTY STANDARD; PRT; 25 AA.  
AC P40725; 25 AA.  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE HYPOTHETICAL PROTEIN IN RARD-PLDA INTERGENIC REGION (FRAGMENT).  
GN YIGI.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.

```
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X76900; -; NOT_ANNOTATED_CDS.
DR StyGene; SG10461; yigI.
KW Hypothetical protein.
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2827 MW; B5226164481A3CAB CRC64;

Query Match 22.0%; Score 20; DB 1; Length 25;
Best Local Similarity 57.1%; Pred. No. 2.9e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 FYLAMPF 12
Db 18 FVYHMPF 24
```

Search completed: May 8, 2001, 15:14:09  
Job time: 530 sec

Result No.	Query No.	Score	Query			DB	ID	Description
			Match	Length				
1	32	35.2	17	4	Q9UCN0	Q9UCN0	homo sapien	
2	29	31.9	17	4	Q16309	Q16309	homo sapien	
3	29	31.9	17	4	Q16310	Q16310	homo sapien	
4	28	30.8	24	6	Q9TX5	Q9TX5	bos taurus	
5	28	30.8	24	14	Q11341	Q11341	molluscum c	
6	27	29.7	17	10	Q9S8Y2	Q9S8Y2	lupinus arb	
7	25	27.5	16	2	Q30985	Q30985	rhodobacter	
8	25	27.5	17	2	Q9Z633	Q9Z633	chlamydia t	
9	25	27.5	24	10	Q96346	Q96346	brassica na	
10	24.5	26.9	23	4	Q16884	Q16884	homo sapien	
11	24	26.4	15	2	Q9R5A0	Q9R5A0	micrococcus	
12	24	26.4	20	2	Q9R896	Q9R896	chlamydia t	
13	24	26.4	20	6	Q9TR72	Q9TR72	ursus arcto	
14	24	26.4	21	2	Q9S1C0	Q9S1C0	porphyromon	
15	24	26.4	23	11	Q9R231	Q9R231	rattus norv	
16	24	26.4	25	10	Q49883	Q49883	lycopersico	
17	24	26.4	25	11	Q9QU22	Q9QU22	mus sp. 50-	
18	23.5	25.8	20	8	Q9S947	Q9S947	saccharomyc	
19	23	25.3	16	14	Q79458	Q79458	human immun	

425302, 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE GC\*1F PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 GN GC\*1F.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=95242701; PubMed=7725672;

RX Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;

RT "Characterization of mutants of the vitamin-D-binding protein/group

RT specific component: GC aborigine (IAI) from Australian aborigines and

RT South African blacks, and 2A9 from south Germany.";

RL Vox Sang. 68:50-54(1995).

DR EMBL; S77129; AAD14249.1; -

FT NON\_TER 17

SQ SEQUENCE 17 AA; 1852 MW; BB2982750293722C CRC64;

Query Match 31.9%; Score 29; DB 4; Length 17;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 10 MPFATPME 17

Db 1 MPDAPTE 8

RESULT 3

Q16310

ID Q16310 PRELIMINARY; PRT; 17 AA.

AC Q16310;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)

DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

DE GC\*2 PROTEIN (FRAGMENT).

GN GC\*2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=95242701; PubMed=7725672;

RX Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;

RT "Characterization of mutants of the vitamin-D-binding protein/group

RT specific component: GC aborigine (IAI) from Australian aborigines and

RT South African blacks, and 2A9 from south Germany.";

RL Vox Sang. 68:50-54(1995).

DR EMBL; S77130; AAD14250.1; -

FT NON\_TER 17

SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 31.9%; Score 29; DB 4; Length 17;

Best Local Similarity 75.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 10 MPFATPME 17

Db 1 MPDAPTE 8

RESULT 4

Q9TQX5

ID Q9TQX5 PRELIMINARY; PRT; 20 AA.

AC Q9TQX5;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE 4-HYDROXYNONENAL METABOLIZING GLUTATHIONE S-TRANSFERASE

DE (EC 2.5.1.18) (FRAGMENT).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX

RN

RP

RA

RX

RT

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE.

RA MEDLINE=97135072; PubMed=8980630;

RX He N.G., Singhal S.S., Chaubey M., Awasthi S., Zimniak P.,

RA Partridge C.A., Awasthi Y.C.;

RT "Purification and characterization of a 4-hydroxynonenal metabolizing

RT glutathione S-transferase isozyme from bovine pulmonary microvessel

RT endothelial cells.";

RL Biochim. Biophys. Acta 1291:182-188(1996).

SQ SEQUENCE 20 AA; 2292 MW; 12BAA99647E94A80 CRC64;

Query Match

Best Local Similarity 71.4%; Score 28; DB 6; Length 20;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 11 PFATPME 17

Db 5 PFKTPKE 11

RESULT 5

O11341

ID O11341 PRELIMINARY; PRT; 24 AA.

AC O11341;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)

DE SIMILAR TO VARIOLA FLR AND VACCINIA DLR (FRAGMENT).

GN BI-44-1.

OS Moluscum contagiosum virus subtype 1 (MCV1).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Molluscipoxvirus.

OX NCBI\_TaxID=10280;

RN [1]

RP SEQUENCE FROM N.A.

RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,

RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;

RL Virus Genes 0:0-0(0).

DR EMBL; U86918; AAB57974.1; -

FT NON\_TER 1

SQ SEQUENCE 24 AA; 2892 MW; 6052DD2609CAFADE CRC64;

Query Match

Best Local Similarity 45.5%; Score 28; DB 14; Length 24;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 LLEFYLPMPFA 13

Db 12 LLQYVVVYVFA 22

RESULT 6

Q9S8Y2

ID Q9S8Y2 PRELIMINARY; PRT; 17 AA.

AC Q9S8Y2;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE L-ASPARAGINASE ISOFORM A (EC 3.5.1.1).

OS Lupinus arboreus (Tree lupine).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

OC Fabales; Fabaceae; Papilionoideae; Lupinus.

OX NCBI\_TaxID=3872;

RN [1]

RP SEQUENCE.

RA Lough T.J., Chang K.S., Carne A., Monk B.C., Reynolds P.H.,

RA Farnden K.J.;

RL Phytochemistry 31:1519-1527(1992).

SQ SEQUENCE 17 AA; 1877 MW; DB7925924D5F46AF CRC64;

Query Match 29.7%; Score 27; DB 10; Length 17;  
 Best Local Similarity 71.4%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Gaps 0;

OY 8 LAMPFAT 14

DB -1 IAMPENT 7

RESULT 7

ID O30985 PRELIMINARY; PRT; 16 AA.  
 AC O30985;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
 DE PUTATIVE AMMONIA TRANSPORTER (FRAGMENT).  
 CN AMTB.  
 OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
 OC Rhodobacter.  
 OX NCBI\_TaxID=1063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16PHC;  
 RA Qian Y., Tabita F.R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF023909; AAC34723.1;  
 FT NON\_TER 16  
 SQ SEQUENCE 16 AA; 1812 MW; EDB859A962D6FEEA CRC64;

Query Match 27.5%; Score 25; DB 2; Length 16;  
 Best Local Similarity 36.4%; Pred. No. 1e+03;  
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRLLEFYLM 11

DB 4 STVLVFFMLP 14

RESULT 8

ID Q9ZG33 PRELIMINARY; PRT; 17 AA.  
 AC Q9ZG33;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DE SUBTILISIN/CHYMOTRYPSIN INHIBITOR (FRAGMENT).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=L2 434B;  
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
 RT "Gene identification of Chlamydia trachomatis by random DNA sequencing."  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF087336; AAD04110.1;  
 FT NON\_TER 17  
 SQ SEQUENCE 17 AA; 2043 MW; 4FEE704EE041E120 CRC64;

Query Match 27.5%; Score 25; DB 2; Length 17;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 7 YLAMPFAT 14

:| | |

DB 2 FLGQPFCT 9

RESULT 9

ID Q96346 PRELIMINARY; PRT; 24 AA.  
 AC Q96346;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE SLLB.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97143881; PubMed=8989888;  
 RA Yu K., Schafer U., Glavin T.L., Gorling D.R., Rothstein S.J.;  
 RT "Molecular characterization of the S locus in two self-incompatible  
 RT Brassica napus lines."  
 RL Plant Cell 8:2369-2380(1996).  
 DR EMBL; U66192; AAB49422.1;  
 SQ SEQUENCE 24 AA; 2986 MW; FCB6AC1803F9D91A CRC64;

Query Match 27.5%; Score 25; DB 10; Length 24;  
 Best Local Similarity 57.1%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 EFYLM 11

DB 9 QFYLYVP 15

RESULT 10

ID Q16884 PRELIMINARY; PRT; 23 AA.  
 AC Q16884;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE APOLIPOPROTEIN B28.1 (FRAGMENT).  
 GN APOB28.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94284731; PubMed=8014581;  
 RA Talmud P.J., Krul E.S., Pessah M., Gay G., Schonfeld G.,  
 RA Humphries S.E., Infante R.;  
 RT "Donor splice mutation generates a lipid-associated apolipoprotein B-  
 RT 27.6 in a patient with homozygous hypobetalipoproteinemia\*;  
 RL J. Lipid Res. 35:468-477(1994).  
 DR EMBL; S73194; AAB31094.1;  
 KW Lipoprotein.  
 FT NON\_TER 1  
 SQ SEQUENCE 23 AA; 2790 MW; B028E11C8BE1F812 CRC64;

Query Match 26.9%; Score 24.5; DB 4; Length 23;  
 Best Local Similarity 43.8%; Pred. No. 1.8e+03;  
 Matches 7; Conservative 2; Mismatches 0; Indels 7; Gaps 1;

OY 1 SRLLE-----FYLA 9

DB 8 ARLEPCKNERYFIA 23

RESULT 11

Q9R5A0  
ID Q9R5A0 PRELIMINARY; PRT; 15 AA.  
AC Q9R5A0;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE DNA TOPOISOMERASE I (FRAGMENT)  
OS Micrococcus luteus (Micrococcus lysodeikticus).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Micrococccineae; Micrococcaceae; Micrococcus.  
OX NCBI\_TaxID=1270;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93249439; PubMed=8387285;  
RA Anderliuzzi D., Pedrini A.M.;  
RT "Structural similarities between M. luteus and E. coli DNA  
topoisomerase I.";  
RL Biochem. Biophys. Res. Commun. 192:657-664(1993).  
SQ SEQUENCE 15 AA; 1602 MW; 285457836F151383 CRC64;

Query Match 26.4%; Score 24; DB 2; Length 15;  
Best Local Similarity 50.0%; Pred. No. 1.4e+03;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 AMPFATPMEA 18  
|:| |:  
Db 2 AVPAQTPVYA 11

RESULT 12  
Q9R896  
ID Q9R896 PRELIMINARY; PRT; 20 AA.  
AC Q9R896;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE STRESS RESPONSE PROTEIN (FRAGMENT).  
GN HYPB.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L2 434B;  
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
RT "Gene identification of Chlamydia trachomatis by random DNA  
sequencing.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF087285; AAD04062.1;  
DR HSSP; P06139; 1DER.  
FT NON\_TER 1  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 2341 MW; 2E19D821CE26567B CRC64;

Query Match 26.4%; Score 24; DB 2; Length 20;  
Best Local Similarity 54.5%; Pred. No. 1.9e+03;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 7 YLAMPFATPME 17  
|:| |:  
Db 6 YLSSYEATNPE 16

RESULT 13  
Q9TR72  
ID Q9TR72 PRELIMINARY; PRT; 20 AA.  
AC Q9TR72;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
DE CYTOCHROME P450(B-1) (FRAGMENT).

OS Ursus arctos (Brown bear) (Grizzly bear).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Ursus.  
OX NCBI\_TaxID=9644;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95260372; PubMed=7741768;  
RA Yamamoto Y., Masuda M., Kazusaka A., Imaoka S., Funae Y., Fujita S.;  
RT "Purification and characterization of a form of cytochrome P450 from  
bear liver microsomes.";  
RL Biochem. Pharmacol. 49:965-970(1995).  
SQ SEQUENCE 20 AA; 2254 MW; 88BE235244C269E2 CRC64;

Query Match 26.4%; Score 24; DB 6; Length 20;  
Best Local Similarity 44.4%; Pred. No. 1.9e+03;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SRLLEFYLA 9  
|:| |:  
Db 12 SQLIPFFLS 20

RESULT 14  
Q9S1C0  
ID Q9S1C0 PRELIMINARY; PRT; 21 AA.  
AC Q9S1C0;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE TRANSPOSASE (FRAGMENT).  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; CFB group; Bacteroidaceae; Porphyromonas.  
OX NCBI\_TaxID=837;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WB3;  
RA Sawada K., Kokeguchi S., Hongyo H., Sawada S., Miyamoto M., Maeda H.,  
RA Nishimura F., Takashiba S., Murayama Y.;  
RT "Identification by subtractive hybridization of a novel insertion  
sequence specific for virulent strains of Porphyromonas gingivalis.";  
RL Infect. Immun. 67:5621-5625(1999).  
DR EMBL; AB011547; BAAB3478.1;  
FT NON\_TER 1  
FT NON\_TER 21  
SQ SEQUENCE 21 AA; 2543 MW; 08D3F838F8B3E6AF CRC64;

Query Match 26.4%; Score 24; DB 2; Length 21;  
Best Local Similarity 45.5%; Pred. No. 2e+03;  
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 YLAMPFATPME 17  
|:| |:  
Db 5 WLNEFFTPPE 15

RESULT 15  
Q9R231  
ID Q9R231 PRELIMINARY; PRT; 23 AA.  
AC Q9R231;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
DE TRANSFORMING GROWTH FACTOR BETA-1 (FRAGMENT).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FISHER;  
RX MEDLINE=98434461; PubMed=9757003;  
RA Yang Y., Mummy M., Romeo D., Wakefield L.M.;

RT \*Identification of the start sites for the 1.9- and 1.4-kb rat  
RT transforming growth factor-beta1 transcripts and their effect on  
RT translational efficiency.\*;  
RL Gene 219:81-89(1998).  
DR EMBL; AF105069; AAD20222.1; -.  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2550 MW; 26AA10E5D88E68F4 CRC64;

Query Match 26.4%; Score 24; DB 11; Length 23;  
Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 2 RLLEFYLPWF 12  
Db . . . . .  
7 RLLPLLLPLPW 17

Search completed: May 8, 2001, 15:16:08  
Job time: 534 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:05:14 ; Search time 62.11 Seconds  
(without alignments)  
5.567 Million cell updates/sec

Title: US-09-165-546A-13  
Perfect score: 88  
Sequence: 1 TVSGNLTIRLTAAADHRQ 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 110741

Minimum DB seq length: 0  
Maximum DB seq length: 25

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCPUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	35.2	25	1 US-08-087-772A-8	Sequence 8, Appli
2	29	33.0	17	6 5185441-3	Patent No. 5185441
3	29	33.0	23	1 US-08-036-555B-123	Sequence 123, App
4	29	33.0	23	1 US-08-469-569-123	Sequence 123, App
5	29	33.0	23	1 US-08-249-322A-123	Sequence 123, App
6	29	33.0	23	1 US-08-469-526A-123	Sequence 123, App
7	29	33.0	23	2 US-08-734-591A-123	Sequence 123, App
8	29	33.0	23	2 US-08-469-660-123	Sequence 123, App
9	29	33.0	23	4 US-08-470-335-123	Sequence 123, App
10	29	33.0	23	4 US-08-735-021-123	Sequence 123, App
11	29	33.0	23	4 US-08-734-664A-123	Sequence 123, App
12	29	33.0	23	5 PCT-US94-05083C-119	Sequence 119, App
13	29	33.0	23	5 PCT-US95-08846A-123	Sequence 123, App
14	29	33.0	24	3 US-08-460-576-7	Sequence 123, App
15	29	33.0	25	1 US-08-029-402-5	Sequence 7, Appli
16	28	31.8	21	1 US-07-593-657-1	Sequence 1, Appli
17	27	30.7	25	2 US-08-726-306A-67	Sequence 67, Appli
18	26	29.5	16	1 US-08-036-555B-121	Sequence 121, App
19	26	29.5	16	1 US-08-469-569-121	Sequence 121, App
20	26	29.5	16	1 US-08-249-322A-121	Sequence 121, App
21	26	29.5	16	1 US-08-469-526A-121	Sequence 121, App
22	26	29.5	16	2 US-08-734-591A-121	Sequence 121, App
23	26	29.5	16	2 US-08-469-660-121	Sequence 121, App
24	26	29.5	16	4 US-08-470-335-121	Sequence 121, App
25	26	29.5	16	4 US-08-735-021-121	Sequence 121, App
26	26	29.5	16	4 US-08-734-664A-121	Sequence 121, App
27	26	29.5	16	5 PCT-US94-05083C-117	Sequence 117, App

28	26	29.5	16	5 PCT-US95-06846A-121	Sequence 121, App
29	26	29.5	18	2 US-08-702-105A-15	Sequence 15, Appl
30	26	29.5	18	3 US-08-702-110A-15	Sequence 15, Appl
31	25	28.4	9	4 US-09-258-754-253	Sequence 253, App
32	25	28.4	11	1 US-08-665-966-4	Sequence 4, Appli
33	25	28.4	11	3 US-09-041-780-4	Sequence 4, Appli
34	25	28.4	19	1 US-08-466-033-79	Sequence 79, Appl
35	25	28.4	19	2 US-08-444-733-79	Sequence 79, Appl
36	25	28.4	19	2 US-08-464-134-79	Sequence 79, Appl
37	25	28.4	19	2 US-08-461-361-79	Sequence 79, Appl
38	25	28.4	19	2 US-08-485-910-79	Sequence 79, Appl
39	25	28.4	19	5 PCT-US95-06266-63	Sequence 63, Appl
40	25	28.4	20	1 US-08-289-653-3	Sequence 3, Appli
41	25	28.4	20	1 US-08-318-193-63	Sequence 63, Appl
42	25	28.4	20	1 US-08-385-186-10	Sequence 10, Appl
43	25	28.4	20	4 US-09-230-421-11	Sequence 11, Appl
44	25	28.4	21	2 US-08-746-283-24	Sequence 24, Appl
45	25	28.4	21	2 US-08-746-257A-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1  
US-08-087-772A-8  
; Sequence 8, Application US/08087772A  
; Patent No. 5691155  
; GENERAL INFORMATION:  
; APPLICANT: Nahmias, Clara  
; APPLICANT: Emorine, Jean L.  
; APPLICANT: Strosberg, Donny A.  
; TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine  
; TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bell, Seltzer, Park & Gibson  
; STREET: Post Office Drawer 34009  
; CITY: Charlotte  
; STATE: No. 5691155th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/087,772A  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Linker, Raymond O.  
; REGISTRATION NUMBER: 26,419  
; REFERENCE/DOCKET NUMBER: 3399-195  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-087-772A-8

Query Match 35.2%; Score 31; DB 1; Length 25;  
Best Local Similarity 66.7%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TVSGNLTIR 9  
||| |||

Db 13 TVGNNLLVI 21

RESULT 2

5185441-3

; Patent No. 5185441

; APPLICANT: WALLNER, BARBARA P.; HESIONS, CATHERINE

; TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA

; MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYMPHOCYTE

; FUNCTION ASSOCIATED ANTIGEN-3

; NUMBER OF SEQUENCES: 41

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/237,309

; FILING DATE: 26-AUG-1988

; SEQ ID NO:3:

; LENGTH: 17

5185441-3

Query Match

Best Local Similarity 33.0%; Score 29; DB 6; Length 17;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 TVSGNLLTIRLTA 13

Db 5 TVSGSLTIYNLTS 17

RESULT 3

US-08-036-555B-123

; Sequence 123, Application US/08036555B

; Patent No. 5530109

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

; APPLICANT: Chen, Maio Su; Hiles, Ian

; TITLE OF INVENTION: Glial Mitogenic Factors, Their

; TITLE OF INVENTION: Preparation and Use

; NUMBER OF SEQUENCES: 184

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/036,555B

; FILING DATE: 24-MAR-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 5250.4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 123:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 23

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

US-08-036-555B-123

Query Match

Best Local Similarity 33.0%; Score 29; DB 1; Length 23;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16

Db 1 LLTVRLGAWGH 11

RESULT 4

US-08-469-569-123

; Sequence 123, Application US/08469569

; Patent No. 5606032

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

; APPLICANT: Chen, Maio Su; Hiles, Ian

; TITLE OF INVENTION: Glial Mitogenic Factors, Their

; TITLE OF INVENTION: Preparation and Use

; NUMBER OF SEQUENCES: 184

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/469,569

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/036,555

; FILING DATE: 24-MAR-1993

; APPLICATION NUMBER: 07/965,173

; FILING DATE: 23-OCT-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/940,389

; FILING DATE: 03-SEP-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/907,138

; FILING DATE: 30-JUN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/863,703

; FILING DATE: 03-APRIL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.K. 91 07566.3

; FILING DATE: 10-APRIL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Tsai, Christine H.

; REGISTRATION NUMBER: 34,266

; REFERENCE/DOCKET NUMBER: LUD 5250.4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 123:

SEQUENCE CHARACTERISTICS:

LENGTH: 23

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-469-569-123

Query Match 33.0%; Score 29; DB 1; Length 23;

Best Local Similarity 54.5%; Pred. No. 47;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16

DB 1 LLTVRLGAWGH 11

RESULT 5

US-08-249-322A-123

Sequence 123, Application US/08249322A

Patent No. 5716930

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew; Stroobant, Paul;

APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;

APPLICANT: Chen, Maio Su; Hiles, Ian

TITLE OF INVENTION: Glial Mitogenic Factors, Their

PREPARATION AND USE

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/249,322A

FILING DATE: 26-MAY-1994

CLASSIFICATION: 435

Prior APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

Prior APPLICATION DATA:

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

Prior APPLICATION DATA:

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

Prior APPLICATION DATA:

APPLICATION NUMBER: 07/907,138

FILING DATE: 30-JUN-1992

Prior APPLICATION DATA:

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APRIL-1992

Prior APPLICATION DATA:

APPLICATION NUMBER: U.K. 91 07566.3

FILING DATE: 10-APRIL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: LUD 250.4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 638-3884

INFORMATION FOR SEQ ID NO: 123:

SEQUENCE CHARACTERISTICS:

LENGTH: 23

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-469-526A-123

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-249-322A-123

Query Match 33.0%; Score 29; DB 1; Length 23;

Best Local Similarity 54.5%; Pred. No. 47;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16

DB 1 LLTVRLGAWGH 11

RESULT 6

US-08-469-526A-123

Sequence 123, Application US/08469526A

Patent No. 5792849

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew

APPLICANT: Stroobant, Paul

APPLICANT: Minghetti, Luisa

APPLICANT: Waterfield, Michael

APPLICANT: Marchionni, Mark

APPLICANT: Chen, Maio Su

APPLICANT: Hiles, Ian

TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

PREPARATION AND USE

NUMBER OF SEQUENCES: 187

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,526A

FILING DATE: 06 June 1995

CLASSIFICATION: 435

Prior APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

APPLICATION NUMBER: 07/907,138

FILING DATE: 03-JUN-1992

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APRIL-1992

APPLICATION NUMBER: U.K. 91 07566.3

FILING DATE: 10-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Bieker-Brady, Kristina

REGISTRATION NUMBER: 39,109

REFERENCE/DOCKET NUMBER: 04585/00200A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 123:

SEQUENCE CHARACTERISTICS:

LENGTH: 23

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-469-526A-123

Query Match 33.0%; Score 29; DB 1; Length 23;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16  
:||:|||||  
Db 1 LLTVRLGAWGH 11

RESULT 7  
US-08-734-591A-123  
; Sequence 123, Application US/08734591A  
; Patent No. 5854220  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Hiles, Ian  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 187  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Ebing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM Compatible Pentium  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: WordPerfect (Version 7.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/734,591A  
; FILING DATE: 22-OCT-1996

CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/470,335  
; FILING DATE: 06-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/036,555  
; FILING DATE: 03-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940,389  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/907,138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863,703  
; FILING DATE: 03-APR-1992

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: UK 91 07566.3  
; FILING DATE: 10-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bleker-Brady, Kristina  
; REGISTRATION NUMBER: 39,109  
; REFERENCE/DOCKET NUMBER: 04585/00200P  
; TELEPHONE: (617) 428-0200  
; TELEFAX: (617) 428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 123:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 23  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-734-591A-123

Query Match 33.0%; Score 29; DB 2; Length 23;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16  
:||:|||||  
Db 1 LLTVRLGAWGH 11

RESULT 8  
US-08-469-660-123  
; Sequence 123, Application US/08469660  
; Patent No. 5876973  
; GENERAL INFORMATION:  
; APPLICANT: Gwynne, David I.; Marchionni, Mark;  
; APPLICANT: McBurney, Robert N.  
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,  
; TITLE OF INVENTION: THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; ZIP: 0211-2804

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,660  
; FILING DATE:

CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/011,396  
; FILING DATE: 29-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/984,085  
; FILING DATE: 01-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/951,747  
; FILING DATE: 25-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/927,337  
; FILING DATE: 10-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 04585/017004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: 200154

; INFORMATION FOR SEQ ID NO: 123:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-469-660-123

Query Match 33.0%; Score 29; DB 2; Length 23;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16  
:||:| | |  
Db 1 LITVRLGAWGH 11

## RESULT 9

US-08-470-335-123  
; Sequence 123, Application US/08470335F  
; Patent No. 6147190  
; GENERAL INFORMATION:  
; APPLICANT: GOODEARL, ANDREW  
; APPLICANT: STROOBANT, PAUL  
; APPLICANT: MINGHETTI, LUISA  
; APPLICANT: WATERFIELD, MICHAEL  
; APPLICANT: MARCHIONNI, MARK  
; APPLICANT: CHEN, MARIO S.  
; APPLICANT: HILES, IAN  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04585/00200B  
; CURRENT APPLICATION NUMBER: US/08/470.335F  
; CURRENT FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036.555  
; EARLIER FILING DATE: 1993-03-24  
; NUMBER OF SEQ ID NOS: 252  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 123  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-470-335-123

Query Match 33.0%; Score 29; DB 4; Length 23;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16  
:||:| | |  
Db 1 LITVRLGAWGH 11

## RESULT 10

US-08-735-021-123  
; Sequence 123, Application US/08735021B  
; Patent No. 6194377  
; GENERAL INFORMATION:  
; APPLICANT: GOODEARL, ANDREW  
; APPLICANT: STROOBANT, PAUL  
; APPLICANT: MINGHETTI, LUISA  
; APPLICANT: WATERFIELD, MICHAEL  
; APPLICANT: MARCHIONNI, MARK  
; APPLICANT: CHEN, MARIO S.  
; APPLICANT: HILES, IAN  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: 04585/00200L  
; CURRENT APPLICATION NUMBER: US/08/735.021B  
; CURRENT FILING DATE: 1996-10-22  
; EARLIER APPLICATION NUMBER: 08/472.065  
; EARLIER FILING DATE: 1995-06-06  
; EARLIER APPLICATION NUMBER: 08/036.555  
; EARLIER FILING DATE: 1993-03-24  
; EARLIER APPLICATION NUMBER: 07/965.173  
; EARLIER FILING DATE: 1992-10-23  
; EARLIER APPLICATION NUMBER: 07/940.389  
; EARLIER FILING DATE: 1992-09-03  
; EARLIER APPLICATION NUMBER: 07/907.138  
; EARLIER FILING DATE: 1992-06-30  
; EARLIER APPLICATION NUMBER: 07/863.703  
; EARLIER FILING DATE: 1992-04-03  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 123  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-08-735-021-123

Query Match 33.0%; Score 29; DB 4; Length 23;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 ILTIRLTAADH 16  
:||:| | |  
Db 1 LITVRLGAWGH 11

## RESULT 11

US-08-734-664A-123  
; Sequence 123, Application US/08734664A  
; Patent No. 6204241  
; GENERAL INFORMATION:  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Stroobant, Paul  
; APPLICANT: Minghetti, Luisa  
; APPLICANT: Waterfield, Michael  
; APPLICANT: Marchionni, Mark  
; APPLICANT: Chen, Mario  
; APPLICANT: Hiles, Ian  
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 187  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM Compatible Pentium  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/734.664A  
; FILING DATE: 22-OCT-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/249.322  
; FILING DATE: 26-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/036.555  
; FILING DATE: 24-MAR-1993  
; APPLICATION DATA:  
; APPLICATION NUMBER: 07/965.173  
; FILING DATE: 23-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/940.389  
; FILING DATE: 03-SEP-1992  
; APPLICATION DATA:  
; APPLICATION NUMBER: 07/907.138  
; FILING DATE: 30-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/863.703  
; FILING DATE: 03-APR-1992  
; APPLICATION DATA:  
; APPLICATION NUMBER: UK 91 07566.3  
; FILING DATE: 10-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bleker-Brady, Kristina  
; REGISTRATION NUMBER: 39,109  
; REFERENCE/DOCKET NUMBER: 04585/00200J  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 428-0200  
 TELEFAX: (617) 428-7045  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 123:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 23  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 US-08-734-664A-123

Query Match 33.0%; Score 29; DB 4; Length 23;  
 Best Local Similarity 54.5%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

OY 6 ILTIRLTAADH 16  
 :||:| | | |  
 Db 1 LLTVRLGAWGH 11

RESULT 12  
 PCT-US94-05083C-119  
 ; Sequence 119, Application PC/TUS9405083C  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robert Sklar, Mark Marchionni,  
 ; APPLICANT: David I. Gwynne  
 ; TITLE OF INVENTION: METHODS FOR ALTERING  
 ; TITLE OF INVENTION: MUSCLE CONDITION  
 ; NUMBER OF SEQUENCES: 185  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360  
 ; MEDIUM TYPE: kb storage  
 ; COMPUTER: IBM  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/05083C  
 ; FILING DATE: 06-MAY-94  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/209,204  
 ; FILING DATE: 08-MAR-94  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/059,022  
 ; FILING DATE: 06-MAY-93  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Clark, Paul T.  
 ; REGISTRATION NUMBER: 30,162  
 ; REFERENCE/DOCKET NUMBER: 04585/028W01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 119:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; PCT-US94-05083C-119

Query Match 33.0%; Score 29; DB 5; Length 23;  
 Best Local Similarity 54.5%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

OY 6 ILTIRLTAADH 16  
 :||:| | | |  
 Db 1 LLTVRLGAWGH 11

RESULT 13  
 PCT-US95-06846A-123  
 ; Sequence 123, Application PC/TUS9506846A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goodearl, Andrew David; Stroobant, Paul;  
 ; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;  
 ; APPLICANT: Chen, Maio Su; Hiles, Ian  
 ; TITLE OF INVENTION: Glial Mitogenic Factors, Their  
 ; TITLE OF INVENTION: Preparation and Use  
 ; NUMBER OF SEQUENCES: 178  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 ; COMPUTER: IBM  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/06846A  
 ; FILING DATE: 25-MAY-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/249,322  
 ; FILING DATE: 26-MAY-1994  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/036,555  
 ; FILING DATE: 24-MAR-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/965,173  
 ; FILING DATE: 23-OCT-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/940,389  
 ; FILING DATE: 03-SEP-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/907,138  
 ; FILING DATE: 30-JUN-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/863,703  
 ; FILING DATE: 03-APRIL-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: U.K. 91 07566.3  
 ; FILING DATE: 10-APRIL-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hanson, Norman D.  
 ; REGISTRATION NUMBER: 30,946  
 ; REFERENCE/DOCKET NUMBER: LUD 5250.5  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; TELEFAX: (212) 838-3884  
 ; INFORMATION FOR SEQ ID NO: 123:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; PCT-US95-06846A-123

Query Match 33.0%; Score 29; DB 5; Length 23;  
 Best Local Similarity 54.5%; Pred. No. 47;  
 Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

OY 6 ILTIRLTAADH 16

Db 1 LITVRLGAWGH 11

## RESULT 14

US-08-460-576-7  
; Sequence 7, Application US/08460576  
; Patent No. 6033903

## ; GENERAL INFORMATION:

; APPLICANT: SISK, William P.

; TITLE OF INVENTION: A METHOD OF EXPRESSING AND SECRETING

; TITLE OF INVENTION: SOLUBLE EXTRACELLULAR DOMAINS OF HUMAN GONADOTROPIN

; TITLE OF INVENTION: HORMONE RECEPTORS

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/460,576

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: YUN, Allen C.

; REGISTRATION NUMBER: 37,971

; REFERENCE/DOCKET NUMBER: SISK-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-628-5197

; TELEFAX: 202-737-3528

; TELEX: 248633

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 24 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-460-576-7

Query Match 33.0%; Score 29; DB 3; Length 24;

Best Local Similarity 46.7%; Pred. No. 49;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 VSGNLTIRLTAAADH 16

|| :|| :|| |

Db 2 VSAIVLVLLAAAH 16

## RESULT 15

US-08-029-402-5

; Sequence 5, Application US/08029402

; Patent No. 5516657

; GENERAL INFORMATION:

; APPLICANT: Murphy, Cheryl I.

; APPLICANT: Young, Elhu

; TITLE OF INVENTION: Baculovirus Vectors For Expression of

; TITLE OF INVENTION: Secretory and Membrane-Bound Proteins

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1225 Connecticut Avenue

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/029,402

; FILING DATE: 19930305

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fox, Samuel L.

; REGISTRATION NUMBER: 30,353

; REFERENCE/DOCKET NUMBER: 0614.0860001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 466-0800

; TELEFAX: (202) 833-8716

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-029-402-5

Query Match 33.0%; Score 29; DB 1; Length 25;

Best Local Similarity 46.7%; Pred. No. 52;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 VSGNLTIRLTAAADH 16

|| :|| :|| |

Db 2 VSAIVLVLLAAAH 16

Search completed: May 8, 2001, 15:05:15

Job time: 111 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2001, 15:14:09 ; Search time 40.07 Seconds  
(without alignments)  
15,388 Million cell updates/sec

Title: US-09-165-546A-13

Perfect score: 88

Sequence: 1 TVSGNLTIRLTAAHRQ 18

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 1423

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	28.4	20	1 CHP_THICU	P80486 thiobacilla
2	24.5	27.8	24	1 LPTN_ECOLI	P09408 escherichia
3	24	27.3	12	1 DCML_PSECA	P19919 pseudomonas
4	23	26.1	21	1 PSBF_SYNVU	P12239 synechococc
5	23	26.1	23	1 PRP1_RAT	P10165 rattus norv
6	22	25.0	20	1 TPX_CLOPA	P81361 clostridium
7	21	23.9	20	1 PGK_CLOPA	P81346 clostridium
8	20	22.7	15	1 CDN2_LITGI	P56247 litorea gil
9	20	22.7	23	1 LPL_CORGL	P42456 corynebacte
10	20	22.7	23	1 PAPI_MANSE	P30253 manduca sex
11	20	22.7	25	1 BGBP_PENVA	P81182 penaeus van
12	20	22.7	25	1 Y15_BPT3	P20835 bacterioph
13	19	21.6	13	1 TEMC_RANTE	P56918 rana tempor
14	19	21.6	13	1 TEMD_RANTE	P56919 rana tempor
15	19	21.6	13	1 FIBA_RANTE	P56920 rana tempor
16	19	21.6	19	1 FIBA_MUNMU	P14457 muntiacus m
17	19	21.6	20	1 FRE3_LITIN	P56249 litorea inf
18	19	21.6	21	1 SOD2_PICAB	P29428 picea abies
19	19	21.6	22	1 HS71_LEITA	P55938 leishmania
20	19	21.6	22	1 NUO7_SOLTU	P80730 solanum tub
21	19	21.6	23	1 PAP2_MANSE	P30254 manduca sex
22	19	21.6	24	1 75CH_PELLU	P15526 pelodictyon
23	19	21.6	24	1 DHAG_COMTE	P80705 comamonas t
24	19	21.6	25	1 FLB3_TREHY	P80161 treponema h
25	18.5	21.0	20	1 UCRO_EQUAR	P81247 equisetum a
26	18	20.5	10	1 FARP_MYTED	P42560 mytilus edu
27	18	20.5	13	1 RL30_SALTY	O54300 salmonella
28	18	20.5	15	1 CDN3_LITGI	P56248 litorea gil
29	18	20.5	15	1 FRE2_LITIN	P82022 litorea inf
30	18	20.5	15	1 UC13_MAIZE	P80619 zea mays (m
31	18	20.5	19	1 ALL7_OLEEU	P81430 olea europ
32	18	20.5	19	1 FIBA_BUBBU	P14442 bubalus bub
33	18	20.5	19	1 UKA1_HUMAN	P31940 homo sapien

#### ALIGNMENTS

##### RESULT 1

CHP\_THICU STANDARD; PRT; 20 AA.  
AC P80486;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE CHEMOTHEROTROPH-SPECIFIC PROTEIN (FRAGMENT).  
OS Thiobacillus cuprinus.  
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas.  
OX NCBI\_TaxID=36860;  
RN [1]  
RP SEQUENCE.  
RC STRAIN-DSM 5494;  
RA Marin I., Amaro A.M., Jerez C.A., Amils R., Abad J.P.;  
RL Submitted (SEP-1995) to the SWISS-PROT data bank.  
CC -!- MISCELLANEOUS: FOUND SPECIFICALLY IN CELLS CULTURED  
CC CHEMOTHEROTROPHICALLY.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2167 MW; E5BAEC1BA3238A0A CRC64;

Query Match 28.4%; Score 25; DB 1; Length 20;  
Best Local Similarity 28.6%; Pred. NO. 2.3e+02;  
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

##### QY 1 TVSGNLTIRLTAA 14

I : I : : : I I

##### DB 3 TYTGTVINQTFPA 16

##### RESULT 2

LPTN\_ECOLI STANDARD; PRT; 24 AA.  
ID LPTN\_ECOLI  
AC P09408;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE TRYPTOPHANASE LEADER PEPTIDE.  
GN TNAL OR TNAC.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=86033634; PubMed=3902796;  
RA Stewart V., Yanofsky C.;  
RT "Evidence for transcription antitermination control of tryptophanase  
operon expression in Escherichia coli K-12.";  
RL J. Bacteriol. 164:731-740(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE=93315143; PubMed=7686882;

```

RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RL genome: organizational symmetry around the origin of replication.";
CC Genomics 16:551-561(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11990; AAA24678.1; -.
DR EMBL; LI0328; AAG62058.1; -.
DR EMBL; AF000448; AAC76730.1; -.
DR EcoGene; EG11276; tnaL.
KW Leader peptide.
SQ SEQUENCE 24 AA; 2894 MW; 65E3987EA7C052F9 CRC64;

Query Match 27.8%; Score 24.5; DB 1; Length 24;
Best Local Similarity 36.4%; Pred. No. 3.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 9; Gaps 1;

QY 5 NLTIRLTA-----ADHR 17
DB 2 NILHICVTSKWFNIDNKIVDHR 23

RESULT 3
DCML_PSECA
ID DCML_PSECA STANDARD; PRT; 12 AA.
AC PL919;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] LARGE CHAIN (EC 1.2.2.4)
DE (FRAGMENT).
OS Pseudomonas carboxydovorans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Oligotropha.
OX NCBI_TaxID=40137;
RN [1]
RP SEQUENCE.
RC STRAIN=OM5;
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of Co dehydrogenase structural genes in
RL carboxydrotrophic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 +
CC 2 H(+) + FERRICCYTOCHROME B-561.
CC -1- COFACTOR: MOLYBDENUM.
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
CC PIR; PLO138; PLO138.
KW Oxidoreductase; Molybdenum.
FT VARIANT 11 12 GE -> EK.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1289 MW; 9B170C688E6B02D1 CRC64;

Query Match 27.3%; Score 24; DB 1; Length 12;
Best Local Similarity 45.5%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 NLTIRLTAAD 15
DB 2 NLTQVETPAGE 12

RESULT 4

```

```

PSBE_SYNVU
ID PSBE_SYNVU STANDARD; PRT; 21 AA.
AC P12239;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE CYTOCHROME B559 BETA CHAIN (FRAGMENT).
GN PSBF.
OS Synecococcus vulcanus.
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32053;
RN [1]
RP SEQUENCE.
RA Ikeuchi M., Koike H., Inoue Y.;
RT "Identification of psbI and psbL gene products in cyanobacterial
RT photosystem II reaction center preparation.";
RL FEBS Lett. 251:155-160(1989).
CC -1- FUNCTION: THIS B-TYPE CYTOCHROME IS TIGHTLY ASSOCIATED WITH THE
CC REACTION CENTER OF PHOTOSYSTEM II AND POSSIBLY IS PART OF THE
CC WATER-OXIDATION COMPLEX.
CC -1- SUBUNIT: DIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SIMILARITY: BELONGS TO THE PSBE / PSBF FAMILY.
DR PIR; S05031; S05031.
DR InterPro; IPR001417; -.
DR PROSITE; PS00537; CYTOCHROME_B559; PARTIAL.
KW Photosystem II; Heme; Electron transport; Transmembrane.
FT DOMAIN 1 18 STROMAL (POTENTIAL).
FT TRANSMEM 19 >21 POTENTIAL.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2331 MW; F22B51A673023CC9 CRC64;

Query Match 26.1%; Score 23; DB 1; Length 21;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 VSGNLTIRLTA 13
DB 10 VSYPIFTVRXVA 21

RESULT 5
PRPI_RAT
ID PRPI_RAT STANDARD; PRT; 23 AA.
AC P10165;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ACIDIC PROLINE-RICH PROTEIN PRP18 PRECURSOR (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RX MEDLINE=86033799; PubMed=3840480;
RA Clements S., Mehansho H., Carlson D.M.;
RT "Novel multigene families encoding highly repetitive peptide
RT sequences. Sequence analyses of rat and mouse proline-rich protein
RT cDNAs.";
RL J. Biol. Chem. 260:13471-13477(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11899; AAA41956.1; -.
KW Repeat; Parotid gland; Multigene family; Saliva; Signal.

```

FT SIGNAL 1 ?  
 FT CHAIN ? >23 ACIDIC PROLINE-RICH PROTEIN PRP18.  
 FT NON\_TER 23  
 SQ SEQUENCE 23 AA; 2380 MW; 875B4F61FD056949 CRC64;

Query Match 26.1%; Score 23; DB 1; Length 23;  
 Best Local Similarity 55.6%; Pred. No. 6e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 ILTIRLTA 14

DB 1 MLVLLTAA 9

# RESULT 6

TPX\_CLOPA STANDARD; PRT; 20 AA.  
 AC P81361;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PROBABLE THIOL PEROXIDASE (EC 1.11.1.-) (CP 42) (FRAGMENT).  
 OS Clostridium pasteurianum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 CC Clostridium.  
 OX NCBI\_TaxID=1501;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=W5;  
 RX MEDLINE=98291870; PubMed=9629918;  
 RA Flengsrud R., Skjeldal L.;  
 RT "Two-dimensional gel electrophoresis separation and N-terminal  
 sequence analysis of proteins from Clostridium pasteurianum W5.";  
 RL Electrophoresis 19:802-806(1998).  
 CC -!- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR  
 H2O2(2) (BY SIMILARITY).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 4.6, ITS MW IS: 20.2 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.  
 DR InterPro: IPR02065;  
 DR PROSITE: PS01265; TPX; PARTIAL.  
 KW Oxidoreductase; Peroxidase.  
 FT NON\_TER 20  
 SQ SEQUENCE 20 AA; 2193 MW; 08178FCD2782E765 CRC64;

Query Match 25.0%; Score 22; DB 1; Length 20;  
 Best Local Similarity 40.0%; Pred. No. 7.7e+02;  
 Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TVSGNLTIR 10

DB 4 TFOGNEVTIQ 13

# RESULT 7

PGK\_CLOPA STANDARD; PRT; 20 AA.  
 AC P81346;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PUTATIVE PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (CP 11) (FRAGMENT).  
 GN PGK.  
 OS Clostridium pasteurianum.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 CC Clostridium.  
 OX NCBI\_TaxID=1501;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=W5;  
 RX MEDLINE=98291870; PubMed=9629918;

RA Flengsrud R., Skjeldal L.;  
 RT "Two-dimensional gel electrophoresis separation and N-terminal  
 sequence analysis of proteins from Clostridium pasteurianum W5.";  
 RL Electrophoresis 19:802-806(1998).  
 CC -!- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE = ADP +  
 3-PHOSPHO-D-GLYCEROYL PHOSPHATE.  
 CC -!- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 5.6, ITS MW IS: 56.2 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.  
 DR InterPro: IPR001576;  
 DR PROSITE: PS00111; PGLYCERATE\_KINASE; PARTIAL.  
 KW Transferase; Kinase; Glycolysis.  
 FT NON\_TER 20  
 SQ SEQUENCE 20 AA; 2355 MW; 749E31E595C85529 CRC64;

Query Match 23.9%; Score 21; DB 1; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 NILTI 9

DB 3 NLLTI 7

# RESULT 8

CDN2\_LITGI STANDARD; PRT; 15 AA.  
 AC P56247;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CAERIDIN 2.  
 OS Litoria gilleni.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
 OC Litoria.  
 OX NCBI\_TaxID=39405;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Parotoid gland;  
 RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;  
 RT "Peptides from Australian frogs. The structures of the caerins and  
 caeridins from Litoria gilleni.";  
 RL J. Chem. Res. 139:937-961(1993).  
 CC -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR  
 ANTIBIOTIC ACTIVITY.  
 CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 GLANDS.  
 CC -!- MASS SPECTROMETRY: MW=1408; METHOD=FAB.  
 KW Amphibian skin; Amidation.  
 FT MOD\_RES 15  
 SQ SEQUENCE 15 AA; 1410 MW; 06F1BBF272550CBF CRC64;

Query Match 22.7%; Score 20; DB 1; Length 15;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSGNIL 7

DB 5 VGNLL 10

# RESULT 9

LPL\_CORGL STANDARD; PRT; 22 AA.  
 ID LPL\_CORGL  
 AC P42456;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

QY	1	TVSGNLTIR	10
	1	:	1
D6	6	TSRANLLLR	15

RESULT 10  
PAP1\_MANSE STANDARD; PRT; 23 AA.  
AC P30253;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE PARALYTIC PEPTIDE I (PP I).  
OS Manduca sexta (tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Hemolymph;  
RX MEDLINE=91302298; PubMed=2071576;  
RA Skinner W.S.; Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,  
RA Quistad G.B.;  
RT "Isolation and identification of paralytic peptides from hemolymph of  
the lepidopteran insects Manduca sexta, Spodoptera exigua, and  
Heliothis virescens";  
RT J. Biol. Chem. 266:12873-12877(1991).  
RL CC -I- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO  
LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE  
HEMOLYPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.  
CC CC  
CC -II- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.  
DR PIR; A39855; A39855.  
KW Hemolymph.  
FT DISULFID  
SQ SEQUENCE 23 AA; 2436 MW; 0B26CB5C29855FE4 CRC64;

```

Query Match          22.7%  Score 20;  DB 1;  Length 25;
Best Local Similarity 30.0%;  Pred. No. 2.2e+03;
Matches 3;  Conservative 5;  Mismatches 2;  Indels 0;  Gaps 0;

QY      1 TVSCNLTIR 10
      ::|||::|
Db       6 SLAGNFSLR 15

RESULT 12
Y15_BPT3
ID      Y15_BPT3      STANDARD;      PRT;      25 AA.
AC      P20835;
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      01-FEB-1991 (Rel. 17, Last annotation update)
DE      HYPOTHETICAL GENE 1.5 PROTEIN.
GN      1-5.
OS      Bacteriophage T3.
OC      Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae;
OC      T7-like phages.
OX      NCBI_TaxID=10759;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LJURIA;
RX      MEDLINE=87226207; PubMed=3586029;
RA      Schmitt M.P., Beck P.J., Kearney C.A., Spence J.L., Digiovanni D.,
RA      Condreay J.P., Molineux I.J.;
RT      "Sequence of a conditionally essential region of bacteriophage T3,
RL      including the primary origin of DNA replication.";
RJ      J. Mol. Biol. 193:479-495(1987).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X17255; CRA35126.1; -  
 DR EMBL; X05031; CAA28701.1; -  
 DR PIR; S09540; S09540.

KW Hypothetical protein.

SQ SEQUENCE 25 AA; 2821 MW; 92B7709153A2950A CRC64;

Query Match 22.7%; Score 20; DB 1; Length 25;

Best Local Similarity 40.0%; Pred. No. 2.2e+03;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 ILTIRLTAAAD 15

: | : | |

Db 9 VATLGLMVAD 18

RESULT 13

TEMP\_RANTE

ID TEMC\_RANTE STANDARD; PRT; 13 AA.

AC P56918;

DT 30-MAY-2000 (Rel. 39, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE TEMPORIN C.

OS Rana temporaria (European common frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.

OX NCBI\_TaxID=8407;

RN [1]

RP SEQUENCE.

RC TISSUE=Skin;

RX MEDLINE=97175050; PubMed=9022710;

RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,

RA Barra D.;

RT "Temporins, antimicrobial peptides from the European red frog Rana

temporaria.";

RL Eur. J. Biochem. 242:788-792(1996).

CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE

CC BACTERIA.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN

CC FAMILY.

KW Amphibian skin; Antibiotic; Amidation; Multigene family.

FT MOD\_RES 13 13

SQ SEQUENCE 13 AA; 1363 MW; 2201403A655B2448 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 13;

Best Local Similarity 50.0%; Pred. No. 1.6e+03;

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSGNIL 7

: | : |

Db 4 ILGNLL 9

RESULT 14

TEMP\_RANTE

ID TEMD\_RANTE STANDARD; PRT; 13 AA.

AC P56919;

DT 30-MAY-2000 (Rel. 39, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE TEMPORIN D.

OS Rana temporaria (European common frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.

OX NCBI\_TaxID=8407;

RN [1]

RP SEQUENCE AND SYNTHESIS.

RC TISSUE=Skin;

RX MEDLINE=97175050; PubMed=9022710;

RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,

RA Barra D.;

RT "Temporins, antimicrobial peptides from the European red frog Rana

temporaria.";

RL Eur. J. Biochem. 242:788-792(1996).

CC -!- FUNCTION: HAS NO ANTIBACTERIAL ACTIVITY.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN

CC FAMILY.

KW Amphibian skin; Amidation; Multigene family.

FT MOD\_RES 13 13

SQ SEQUENCE 13 AA; 1379 MW; 3EF35DFA655B2448 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 13;

Best Local Similarity 50.0%; Pred. No. 1.6e+03;

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSGNIL 7

: | : |

Db 4 IVGNLL 9

RESULT 15

TEMP\_RANTE

ID TEME\_RANTE STANDARD; PRT; 13 AA.

AC P56920;

DT 30-MAY-2000 (Rel. 39, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE TEMPORIN E.

OS Rana temporaria (European common frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.

OX NCBI\_TaxID=8407;

RN [1]

RP SEQUENCE.

RC TISSUE=Skin;

RX MEDLINE=97175050; PubMed=9022710;

RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,

RA Barra D.;

RT "Temporins, antimicrobial peptides from the European red frog Rana

temporaria.";

RL Eur. J. Biochem. 242:788-792(1996).

CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE

CC BACTERIA.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN

CC FAMILY.

KW Amphibian skin; Antibiotic; Amidation; Multigene family.

FT MOD\_RES 13 13

SQ SEQUENCE 13 AA; 1379 MW; 26505DFA79A92448 CRC64;

Query Match 21.6%; Score 19; DB 1; Length 13;

Best Local Similarity 50.0%; Pred. No. 1.6e+03;

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VSGNIL 7

: | : |

Db 4 IIGNLL 9

Search completed: May 8, 2001, 15:14:10

Job time: 531 sec





OS Proteus vulgaris.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Proteus.  
 OX NCBI\_TaxID=585;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94298804; PubMed=8026480;  
 RA Trautwein T., Krauss F., Lottspeich F., Simon H.;  
 RT "The (2R)-hydroxycarboxylate-viologen-oxidoreductase from *Proteus*  
 RT *vulgaris* is a molybdenum-containing iron-sulphur protein.";  
 RL Eur. J. Biochem. 222:1025-1032(1994).  
 SQ SEQUENCE 20 AA; 2145 MW; D12A6GFC51207C63 CRC64;

Query Match 37.5%; Score 33; DB 2; Length 20;  
 Best Local Similarity 70.0%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 SGNLTIRLT 12  
 Db 6 TGNILRLNT 15  
 :|||||  
 :|||||

RESULT 3  
 OQV28 PRELIMINARY; PRT; 20 AA.  
 AC Q9QV28  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE PROCARBOXYPEPTIDASE A1 (EC 3.4.17.1) (FRAGMENT).  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94259059; PubMed=8200353;  
 RA Oppezzo O., Ventura S., Bergman T., Vendrell J., Jornvall H.,  
 RA Aviles F.X.;  
 RT "Procarboxypeptidase in rat pancreas. Overall characterization and  
 RT comparison of the activation processes";  
 RL Eur. J. Biochem. 222:55-63(1994).  
 DR HSP; P09954; iPCA.  
 SQ SEQUENCE 20 AA; 2197 MW; 25E79553BC03E21C CRC64;

Query Match 31.8%; Score 28; DB 11; Length 20;  
 Best Local Similarity 50.0%; Pred. No. 31e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 IRLTAADHRQ 18  
 Db 10 LRLSADEAQ 19  
 :|||:  
 :|||:

RESULT 4  
 Q53503 PRELIMINARY; PRT; 18 AA.  
 AC Q53503  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE UNKNOWN PROTEIN (FRAGMENT).  
 OS Lactobacillus paracasei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 OX NCBI\_TaxID=1597;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95219758; PubMed=7704831;  
 RA Djordjevic G., Bojovic B., Benina A., Topisirovic L.;  
 RT "Cloning of promoter-like sequences from *Lactobacillus paracasei*

RT Subsp. paracasei CGL1 and their expression in *Escherichia coli*,  
 RT *Lactococcus lactis*, and *Lactobacillus reuteri*.";  
 RL Can. J. Microbiol. 40:1043-1050(1994).  
 DR EMBL; S76789; AAB33948.1; -.  
 FT NON\_TER 18  
 SQ SEQUENCE 18 AA; 2113 MW; 4D8930982275619F CRC64;

Query Match 30.7%; Score 27; DB 2; Length 18;  
 Best Local Similarity 45.5%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 LTRLTAAADHR 17  
 Db 1 MTLALTQSDFR 11  
 :|:|:|:|  
 :|:|:|:|

RESULT 5  
 O54460 PRELIMINARY; PRT; 24 AA.  
 ID O54460  
 AC O54460;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 2.7 KDA PROTEIN (FRAGMENT).  
 OS *Erwinia herbicola*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Erwinia*.  
 OX NCBI\_TaxID=557;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98083064; PubMed=9422601;  
 RA Subramanian P.S., Xie G., Xia T., Jensen R.A.;  
 RT "Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate  
 RT synthase from *Neisseria gonorrhoeae* in the context of its membership  
 RT in a protein family containing a subset of 3-deoxy-D-arabino-  
 RT heptulosonate 7-phosphate synthases";  
 RL J. Bacteriol. 180:119-127(1998).  
 DR EMBL; U93355; AAB96401.1; -.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 24 AA; 2725 MW; 34989DD773855CEE CRC64;

Query Match 30.7%; Score 27; DB 2; Length 24;  
 Best Local Similarity 71.4%; Pred. No. 5.7e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 12 TAAADHRQ 18  
 Db 3 TAAEVRQ 9  
 |||:|  
 |||:|

RESULT 6  
 Q9QV49 PRELIMINARY; PRT; 25 AA.  
 ID Q9QV49  
 AC Q9QV49;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE UBIQUITIN HOMOLOG (FRAGMENT).  
 OS *Rattus* sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94063801; PubMed=8244287;  
 RA Wang B., Harwig S.S., Lehrer R.I.;  
 RT "[Rat bladder ubiquitin-like molecule: isolation, purification and N-  
 RT terminal sequencing]";  
 RL J. West China Univ. Med. Sci 24:127-130(1993).  
 DR HSP; P02248; IUBI.

DR	INTERPRO; IPR000626; -	Query Match	29.5%;	Score 26;	DB 6;	Length 23;	
DR	PFAM; PF00240; ubiquitin; 1.	Best Local Similarity	38.5%;	Pred. No. 8.2e+02;			
SQ	SEQUENCE 25 AA; 2808 MW; E22C0441310AA50D CRC64;	Matches	5;	Conservative	3;	Mismatches	0; Gaps 0;
QY	1 TVSGNLTIRLTAAD 15	QY	2 VSGNLTIRLTA 14				
Db	7 TLTKTITLVEPSD 21	Db	7 ISGNVCPDRSTGS 19				
RESULT	7	RESULT	9				
ID	Q53541	ID	Q9UWH9	PRELIMINARY;	PRT;	15 AA.	
AC	Q53541;	AC	Q9UWH9;				
DT	01-NOV-1996 (TREMBLrel. 01, Created)	DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JAN-1999 (TREMBLrel. 09, Last annotation update)	DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)				
DE	ORF2 PROTEIN (FRAGMENT).	DE	ALDEHYDE FERREDOXIN OXIDOREDUCTASE (FRAGMENT).				
OS	Bacillus sp.	OS	Thermococcus sp. (strain 90N-7).				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;	OC	Archaea; Euryarchaeota; Thermococcales; Thermococcaceae.				
OC	Bacillus/Staphylococcus group; Bacillus.	OC	NCBI_TaxID=1409;				
OX	NCBI_TaxID=1409;	OX	NCBI_TaxID=2263;				
RN	[1]	RN	[1]				
RP	SEQUENCE FROM N.A.	RP	SEQUENCE				
RX	MEDLINE=95400017; PubMed=7670201;	RX	MEDLINE=95370155; PubMed=7642503;				
RA	Seto Y., Hashimoto M., Usami R., Hamamoto T., Kudo T., Horikoshi K.;	RA	Heider J., Ma K., Adams M.W.;				
RT	"Characterization of a mutation responsible for an alkali-sensitive	RT	"Purification, characterization, and metabolic function of tungsten-				
RT	mutant, 18224, of alkaliphilic Bacillus sp. strain C-125.";	RT	containing aldehyde ferredoxin oxidoreductase from the				
RL	Biosci. Biotechnol. Biochem. 59:1364-1366(1995).	RT	hyperthermophilic and proteolytic archaeon Thermococcus strain ES-1.";				
RL	EMBL; S79441; AAB35256.1; -	RL	J. Bacteriol. 177:4757-4764(1995).				
FT	NON_TER 1	RL	J. Bacteriol. 177:4757-4764(1995).				
SQ	SEQUENCE 15 AA; 1529 MW; 5BAD7F1318F3B3C CRC64;	SQ	SEQUENCE 15 AA; 1750 MW; 0441670278198619 CRC64;				
Query Match	29.5%;	Query Match	28.4%;	Score 25;	DB 1;	Length 15;	
Best Local Similarity	58.3%;	Best Local Similarity	55.6%;	Pred. No. 7.9e+02;			
Matches	7;	Matches	5;	Conservative	1;	Mismatches	3; Indels 0; Gaps 0;
QY	4 GNLTIRLTAAD 15	QY	4 GNLTIRLTA 12				
Db	1 GIALTITLTAIE 12	Db	6 GKILRVNLT 14				
RESULT	8	RESULT	10				
ID	Q28016	ID	Q9UWL8	PRELIMINARY;	PRT;	19 AA.	
AC	Q28016;	AC	Q9UWL8;				
DT	01-NOV-1996 (TREMBLrel. 01, Created)	DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-NOV-1998 (TREMBLrel. 08, Last annotation update)	DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)				
DE	INSULIN-LIKE GROWTH FACTOR-1 RECEPTOR (FRAGMENT).	DE	UBIQUITIN (FRAGMENT).				
OS	Bos taurus (Bovine).	OS	Thermoplasma acidophilum.				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	OC	Thermoplasma				
OC	Bovidae; Bovinae; Bos.	OC	NCBI_TaxID=2303;				
OX	NCBI_TaxID=9913;	OX	[1]				
RN	[1]	RN	SEQUENCE				
RP	SEQUENCE FROM N.A.	RP	SEQUENCE				
RX	MEDLINE=96432479; PubMed=8835545;	RX	MEDLINE=93314814; PubMed=8392003;				
RA	Moody D.E., Pomp D., Barendse W.;	RA	Wolf S., Lottspeich F., Baumeister W.;				
RT	"Linkage mapping of the bovine insulin-like growth factor-1 receptor	RT	"Ubiquitin found in the archaeobacterium Thermoplasma acidophilum.";				
RT	gene.";	RT	FEBS Lett. 326:42-44(1993).				
RL	Mamm. Genome 7:168-169(1996).	RL	HSSP; P02248; IUBI.				
RL	EMBL; U33122; AAB52601.1; -	DR	INTERPRO; IPR000626; -				
FT	NON_TER 1	DR	PFAM; PF00240; ubiquitin; 1.				
FT	NON_TER 23	DR	PFAM; PF00240; ubiquitin; 1.				
SQ	SEQUENCE 23 AA; 2458 MW; D9814EC8DB0FE853 CRC64;	SQ	SEQUENCE 19 AA; 2123 MW; 5F8DE7501B6D4599 CRC64;				
Query Match	29.5%;	Query Match	28.4%;	Score 25;	DB 1;	Length 19;	
Best Local Similarity	58.3%;	Best Local Similarity	30.8%;	Pred. No. 1e+03;			
Matches	7;	Matches	4;	Conservative	5;	Mismatches	4; Indels 0; Gaps 0;
QY	1 TVSGNLTIRLTA 13	QY	1 TVSGNLTIRLTA 13				
Db	7 TLTKTITLVEVEA 19	Db	7 TLTKTITLVEVEA 19				

```
RESULT 11
Q9PXH2 PRELIMINARY; PRT; 25 AA.
AC Q9PXH2
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE VARIANT TRANSACTIVATION FACTOR (FRAGMENT).
OS Human T-cell lymphotropic virus type 1.
OC Viruses; Retrod viruses; Retroviridae; BLV-HTLV retroviruses.
OX NCBI_TaxID=11908;
RN [1]
RN MEDLINE=95191051; PubMed=7533860;
RA Niewiesk S., Daenke S., Parker C.E., Taylor G., Weber J.,
RA Nightingale S., Bangham C.R.;
RT "Naturally occurring variants of human T-cell leukemia virus type I
RT tax protein impair its recognition by cytotoxic T lymphocytes and the
RT transactivation function of tax.";
RL J. Virol. 69:2649-2653(1995).
SQ SEQUENCE 25 AA; 2638 MW; CF0FBDA4DBB51D26 CRC64;

Query Match 28.4%; Score 25; DB 14; Length 25;
Best Local Similarity 43.8%; Pred. No. 1.3e+03;
Matches 7; Conservative 2; Mismatches 3; Indels 4; Gaps 1;

QY 2 VSGNLTIRLTAAADR 17
Db 11 VSGGLCSARL----HR 22

RESULT 12
Q05694 PRELIMINARY; PRT; 15 AA.
AC Q05694
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE PROLINE-RICH PROTEIN MP6 (FRAGMENT).
GN MP6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN MEDLINE=10090;
RC STRAIN=BALB/C;
RP SEQUENCE FROM N.A.
RX MEDLINE=92020206; PubMed=1747160;
RA Roberts S.G.E., Layfield R., McDonald C.J.;
RT "The mouse proline-rich protein MP6 promoter binds isoprenaline-
RT inducible parotid nuclear proteins via a highly conserved NFkB/rel-
RT like site.";
RL Nucleic Acids Res. 19:5205-5211(1991).
DR EMBL: X61126; CAA43438.1;
FT NON_TER 15
SQ SEQUENCE 15 AA; 1557 MW; C0FF9514A8F70EE4 CRC64;

Query Match 27.3%; Score 24; DB 11; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 ILTIRLTAA 14
Db 1 MLVILLTAA 9

RESULT 13
Q9QVK0 PRELIMINARY; PRT; 19 AA.
AC Q9QVK0
DT 01-MAY-2000 (TReMBLrel. 13, Created)
```

```
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE MEPRIN-A PEPTIDE A2.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RN MEDLINE=91373354; PubMed=1894622;
RA Kounnas M.Z., Wolz R.L., Gorbea C.M., Bond J.S.;
RT "Meprin-A and -B. Cell surface endopeptidases of the mouse kidney.";
RL J. Biol. Chem. 266:17350-17357(1991).
SQ SEQUENCE 19 AA; 1943 MW; DD088D1AB5B71051 CRC64;

Query Match 27.3%; Score 24; DB 11; Length 19;
Best Local Similarity 42.9%; Pred. No. 1.5e+03;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 VSGNLT 8
Db 12 ISGSVIT 18

RESULT 14
Q9R4N1 PRELIMINARY; PRT; 20 AA.
AC Q9R4N1
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CHLOROMUCONATE CYCLOISOMERASE (EC 5.5.1.7) (FRAGMENT).
OS Rhodococcus erythropolis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1833;
RN [1]
RN MEDLINE=95270600; PubMed=7751292;
RA Solyanikova I.P., Maltseva O.V., Vollmer M.D., Golovleva L.A.,
RA Schlomann M.;
RT "Characterization of muconate and chloromuconate cycloisomerase from
RT Rhodococcus erythropolis ICP: indications for functionally convergent
RT evolution among bacterial cycloisomerases.";
RL J. Bacteriol. 177:2821-2826(1995).
SQ SEQUENCE 20 AA; 2163 MW; 7F4098F934BEF72F CRC64;

Query Match 27.3%; Score 24; DB 2; Length 20;
Best Local Similarity 46.7%; Pred. No. 1.6e+03;
Matches 7; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 TVSGNLTIRLTAAAD 15
Db 4 TVSG-----VRTTIVD 14

RESULT 15
Q9TRM1 PRELIMINARY; PRT; 20 AA.
AC Q9TRM1
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE 33 KDA CA(2+)-DEPENDENT CARBOHYDRATE-BINDING PROTEIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN SEQUENCE.
```

RX MEDLINE=93015942; PubMed=1400371;  
RA Kojima K., Ogawa H.K., Seno N., Yamamoto K., Irimura T., Osawa T.,  
RA Matsumoto I.,  
RT "Carbohydrate-binding proteins in bovine kidney have consensus amino  
RL acid sequences of annexin family proteins.";  
J. Biol. Chem. 267:20536-20539(1992).  
DR HSP; P13214; 1ANN.  
SQ SEQUENCE 20 AA; 2107 MW; 45BA0E8B451A5C92 CRC64;

Query Match 27.3%; Score 24; DB 6; Length 20;  
Best Local Similarity 55.6%; Pred. No. 1.6e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 5 NILTIRLTA 13  
|:| | | |  
Db 11 NVLAYRSTA 19

Search completed: May 8, 2001, 15:16:09  
Job time: 535 sec